大学

数学系

自学丛书

概率论与数理统计

GAILÜLUN YU SHULI TONGJI

大学数学系自学丛书

概率论与数理统计

东北师范大学

王铭文　主编

辽宁人民出版社

一九八三年·沈阳

大学数学系自学丛书

概率论与数理统计

王铭文 主编

辽宁人民出版社出版　　辽宁省新华书店发行

（沈阳市南京街6段1里2号）　　朝阳六六七厂印刷

字数：430,000　　开本：850×1168 1/32　　印张：16 3/4

印数：1—14,000

1983年6月第1版　　1983年6月第1次印刷

责任编辑：于　艺　　封面设计：安今生

统一书号：7090·234　　定价：2.10 元

目　录

引　　言

概率论与数理统计是研究随机现象的统计规律性的数学学科．下面加以具体说明．

（一）随机现象

在实际中，我们经常碰到这样一类现象：在同样条件下，多次进行同一试验，或多次观测同一现象，所得结果并不完全一样，面往往有些差异，而且在每次试验或观测之前不能确切预料将出现什么结果。这样的现象叫做随机现象。例如

1．一个灯泡厂所生产的同一型号的灯泡，点燃时数却不完全一样，有的点燃1500小时以上，有的却不到1000小时就坏了。

2．用同一种药物对大豆食心虫进行防治，结果每株大豆的虫食情况却不完全一样，有的虫食率不到2％，有的虫食率达到5％以上。

3．一名炮手，使用同一门炮，按同样射击条件（初始速度v、发射角θ与弹道系数c）多次进行射击，结果炮弹并不完全落在同一点上，有的偏左，有的偏右，有的偏远，有的偏近，散布在一定范围里。

为什么在相同的基本条件下，试验或观测的结果会不完全一样呢？这是因为除了基本条件组外，客观上还存在着很多变化着的次要因素，它们中的每一个对试验或观测的结果的影响都很小，但是它们的变化总起来就使得试验或观测的结果产生差异．比如，在大炮射击时，炮弹飞行的气象条件（如风力、气压、温度、湿度等），弹药的成份以及射手的心理状态等都不可能完全一样．虽然它们中

的每一个对射击结果的影响都是很微小的，但是这些为数众多的次要因素合起来就使弹着点产生偶然偏差。

随着科学技术的发展，对试验结果的精确程度要求愈来愈高，自然应当设法对某些次要因素也加以控制，使试验的结果尽可能的精确，这是科学技术工作的一个重要方面。但是，只着眼于这一方面还不行，因为影响试验结果的次要因素无限多，不可能全部加以控制，所以产生偶然偏差是不可避免的，因此，就应当研究这种偶然变化的规律，并运用这种规律去指导实践。

（二）随机现象的统计规律性

随机现象有没有规律性呢？从表面上看，随机现象似乎是一种没有规律性的现象，其实不然。对随机现象只作个别的观测，一般看不出明显的规律性，但是，对随机现象作大量的观测时，就发现它有某种明显的规律性。比如，对某一目标进行射击，当射击次数不多时，从几个零星的弹着点，看不出什么规律，但当多次射击时，就会看到弹着点的分布呈现出某种规律性：弹着点差不多关于目标中心对称；越靠近中心弹着点越密等等。射击次数越多，这种规律性就越明显。

如果射击目标比炮弹散布的区域大，显然，炮弹的随机散布关系不大，但若炮弹散布的区域比目标大，这时某些炮弹就不能命中目标。于是自然产生了一系列的问题，比如射击的命中率是多少？为了摧毁目标需要多少发炮弹？为了降低炮弹损耗需要采取什么办法？等等。只有掌握了随机现象的统计规律性以后，这些问题才能解决。

第一章　事件与概率

事件及事件的概率是概率论中两个基本的概念，研究事件的关系和运算、概率的性质及概率的计算是概率论中的一些基本问题．这些基本概念及基本问题就是第一章的主要内容．

§1.1　事　　件

（一）基本事件空间

由前所述可知，任一随机现象都与某一试验或观测相联系，为了叙述方便，今后把“试验”、“观测”统一叫做试验．一个试验，如果不能事先准确地预言它的结果，而且在相同条件下可以重复进行，则称其为随机试验．概率论中所说的试验都是指随机试验．对于试验，我们关心的是它的结果，在概率论中，我们把试验的每一个结果叫做一个事件．

比如，为了知道一批大豆的虫食情况，随机从中取出100粒来检查，结果可能是“没有虫豆”，可能是“有一粒虫豆”，可能是“有两粒虫豆”，等等．其中每一个结果都是一个事件．此外，“虫豆数不多于5”，“虫豆数在5与10中间”，“虫豆数在20以上”，等等结果也都是事件．

在上述这些事件中，我们把“没有虫豆”，“有一粒虫豆”，…，“有100粒虫豆”，叫做基本事件；而把“虫豆数不多于5”，“虫豆数在5与10中间”等叫做复合事件．每一个复合事件都是由一些基本事件所组成，如复合事件“虫豆数小于2”，就是由“虫

豆数是0”和“虫豆数是1”两个基本事件所组成．概括说就是，基本事件为不可分解的事件，复合事件为可分解的事件．

所有基本事件组成的集合叫做基本事件空间（也叫样本空间），我们用Ω表示，基本事件（也叫样本点），我们用ω表示．

在上面的例子中，还有两个特殊的事件．一个是“虫豆数在0与100中间”，这是必然的，把它叫做必然事件．一般来说，在试验条件下必然发生的结果叫做必然事件．另一个是“虫豆数不在0与100中间”，这是不可能的，把它叫做不可能事件．一般来说，在试验条件下不可能发生的结果叫做不可能事件．在概率论中主要讨论随机事件，即在试验条件下可能发生也可能不发生的事件．由于偶然与必然是辩证的统一，所以在研究随机现象时不能不涉及必然事件与不可能事件．

对于一个试验，首先应当弄清楚与它对应的基本事件空间是什么，即一切基本事件都是哪些，这是我们弄清事件这一概念的基础．下面举几个例子．

1．在一个口袋中放有红、白、黑三个球，从中随机取一个，观察它的颜色．在这个试验中，基本事件空间Ω由下面三个基本事件所组成：

ω_1=（是红球），ω_2=（是白球），ω_3=（是黑球）．

$$\Omega=\{\omega_1,\omega_2,\omega_3\}$$

2．从放有红、白、黑三个球的口袋中，随机取出一个球，然后放回，再取出一个，观察所取两球的颜色．在这个试验中如果考虑所取两球的顺序，则基本事件空间Ω由下面9个基本事件所组成：

ω_1=（红，红），ω_2=（红，白），ω_3=（红，黑），

ω_4=（白，红），ω_5=（白，白），ω_6=（白，黑），

ω_7=（黑，红），ω_8=（黑，白），ω_9=（黑，黑），

$$\Omega=\{\omega_1,\ \omega_2\cdots,\ \omega_9\}$$

注意，这里不能认为（一黑，一白），是一个基本事件，因为

它是由ω_6＝（白，黑）和ω_8＝（黑，白）两个基本事件所组成的复合事件。

3．观察某电话交换台在某一段时间内接到的呼唤次数，这时基本事件空间由下面可列个基本事件所组成：

$$\omega_i=(\text{接到 } i \text{ 次呼唤}),\quad i=0,1,2,\cdots$$

$$\Omega=\{\omega_0,\omega_1,\cdots\}$$

在以上三例中，基本事件空间都是离散的，即由有限个或可列个基本事件所组成．实际中还有一些试验，它所对应的基本事件空间并不是离散的，而是连续的。比如

4．观察某地区某种小麦的平均亩产，这时基本事件空间由下面的基本事件所组成：

$$\omega_a=(\text{平均亩产是 } a \text{ 斤}),\quad 0\leqslant a<\infty$$

$$\Omega=\{\omega_a,\ 0\leqslant a<\infty\}$$

（二）事件

上面已讲过，任一事件都是由若干个基本事件所组成，所以，实际上事件就是基本事件的某一集合，即基本事件空间的一个子集．今后用大写字母A，B，$C\cdots$表示事件．在上面的例3中，若A表示事件“来到的呼唤数不多于10”，则$A=\{\omega_0,\omega_1,\cdots,\omega_{10}\}$．若$B$表示事件“来到的呼唤数在5与12中间”，则$B=\{\omega_5,\omega_6,\cdots,\omega_{12}\}$．在这里，事件$A$与$\omega_0,\omega_1,\cdots,\omega_{10}$的关系显然是，若$A$发生，则$\omega_0,\omega_1,\cdots,\omega_{10}$中必有一个发生；反之，若$\omega_0,\omega_1,\cdots,\omega_{10}$中的某一个发生，则$A$必发生．简言之，即当且仅当事件$A$所含的一个基本事件发生时$A$发生。

基本事件ω就是单点集$\{\omega\}$，必然事件就是全空间Ω，因为所谓必然事件就是不论Ω中哪一个基本事件发生它都发生，有时用U表示，不可能事件就是空集ϕ，因为所谓不可能事件就是Ω中没有一个基本事件发生能使它发生，有时用V表示。

§1.2 事件的关系和运算

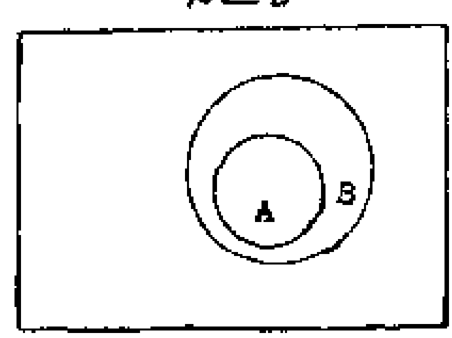
研究一个随机现象，常常同时涉及许多事件，而这些事件是有关系的。了解事件间的关系，能使我们更进一步了解事件的本质，可以从一个事件与其他事件的关系上来了解和掌握这个事件，能使我们通过对简单事件的了解，去研究与其有关的较复杂的事件，这一点在研究随机现象的规律上是十分重要的。

关于事件间的关系和运算，为了直观起见，我们结合下面的试验来说明：向平面上某一矩形内随机掷一点，观察点落的位置。显然，试验的每一结果对应矩形内的一个点，基本事件空间对应矩形内的全部点。

（一）包含

设事件 A = “点落在小圆里”，事件 B = “点落在大圆里”，如图（1,2,1）。显然，若所掷的点落在小圆里，则该点必落在大圆里。也就是说，若 A 发生，则 B 必发生。对此我们说 A 含于 B，或 B 包含 A。一般地有如下定义。

$A\subset B$



图（1.2.1）

定义 1.2.1　如果事件 A 发生必然导致事件 B 发生，则说 B 包含 A，或说 A 含于 B，记为

$$A\subset B$$

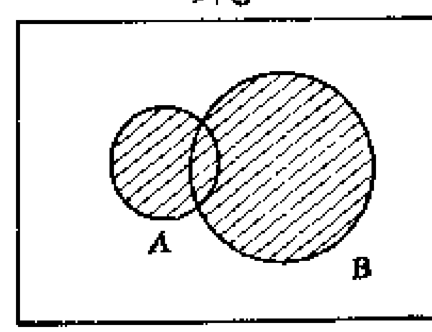
如果 $A\subset B$, $B\subset A$ 同时成立，则说 A 与 B 相等，记为 $A=B$。

由于事件是基本事件空间的子集，所以，事件 A 含于 B 就等价于事件 A 中的每一个基本事件都包含于 B 中。这与集合的包含关系完全一致。因此，在我们这个掷点的例子中，可以直接用小圆表示 A，用大圆表示 B，这样看起来就更直观了。

（二）并（或和）

设事件A，B如(一)，不过这时小圆与大圆的位置如图(1.2.2)．

考虑事件“点落在阴影部分里”．显然，当且仅当点落在两圆中至少一个里时，点落在阴影部分里，也就是说，当且仅当事件A、B中至少一个发生时，该事件发生．这时我们说该事件为事件A与事件B的并．



图（1.2.2）

定义 1.2.2　事件A与事件B至少一个发生，是一个事件，叫做事件A与事件B的并，记为 $A\cup B$．即

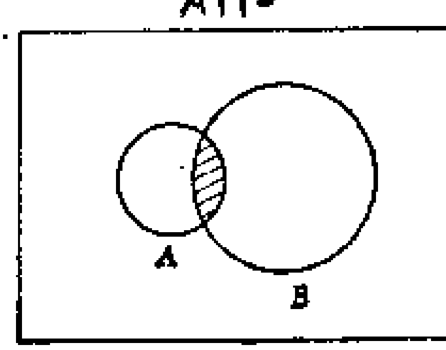
$$A\cup B=\text{“}A\text{与}B\text{至少发生其一”}$$

$A\cup B$ 是由A与B中的所有基本事件所组成的集合．

（三）交（或积）

设事件A、B如（一），而大、小圆位置如图（1.2.3）．

考虑事件“点落在两圆的共同部分里”．显然，当且仅当事件A，B同时发生时，该事件发生．这时我们说该事件是事件A与事件B的交．



图（1.2.3）

定义 1.2.3　事件A与事件B同时发生，是一个事件，叫做事件A与事件B的交，记为$A\cap B$(或AB，或$A\cdot B$)，即

$$A\cap B=\text{“}A\text{与}B\text{同时发生”}$$

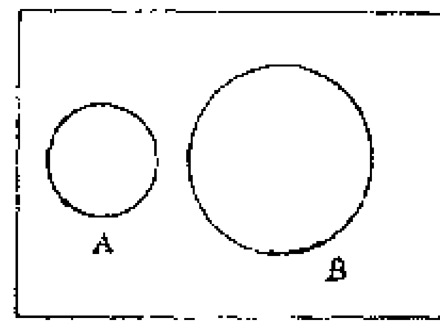
$A\cap B$ 是由同时属于A与B的基本事件所组成．

并与交的概念，可推广到事件为有限个以至可列个的情形中去．若A发生等价于A_1，A_2，…中至少有一个发生，则称A是A_1，A_2，…的并，用 $A=\bigcup\limits_{n=1}^{\infty}A_n$ 表示；若A发生等价于 A_1，A_2，…

同时发生，则称A是A_1，A_2…的交，用 $A=\bigcap_{n=1}^{\infty}A_n$表示。

（四）互斥（或互不相容）

设事件A，B如（一），而大、小圆位置如图（1.2.4），显然，点不能同时落在两个圆里，也就是说A，B不能同时发生。这时我们说A与B互斥。



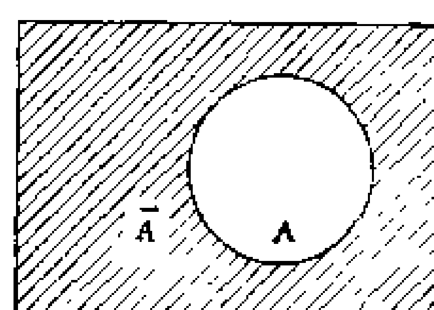
图（1.2.4）

定义 1.2.4　若事件A与事件B不能同时发生，即 $A\cap B=V$，则称A与B互斥，或互不相容。

两个互斥事件不包含共同的基本事件。今后对互斥事件A，B，可记它们的并为 $A+B$。

（五）互逆（或对立）

设事件$A=$“点落在圆内”，考虑事件“点落在圆外”，该事件与A不能同时发生而两者又必发生其一。这时我们说，该事件与A互逆。



图（1.2.5）

定义1.2.5　对于事件A，称满足$A\cup\overline{A}=U$，$A\cap\overline{A}=V$的事件$\overline{A}$为事件A的逆事件（或对立事件）。

显然，A也是$\overline{A}$的逆事件，即$\overline{\overline{A}}=A$，所以，我们说$A$与$\overline{A}$互逆。

$\overline{A}$是由基本事件空间中不包含在A中的基本事件所组成。

注意：互逆必互斥，但互斥未必互逆。比如，事件“击中10环”与“击中9环”是互斥的，但不是互逆的，因为不能说“没有击中10环”就是“击中9环”。击中10环”与“没击中10环”是互逆的。

（六）差

设事件A，B如(一)，大、小圆位置及阴影部分如图(1.2.6)。

考虑事件“点落在阴影部分里”，显然，当且仅当点落在小圆里而同时又不落在大圆里时，点落在阴影部分里，也就是说，当且仅当事件A发生，而事件B不发生时，该事件发生。这时我们说，该事件是事件A与事件B的差。

$A-B$

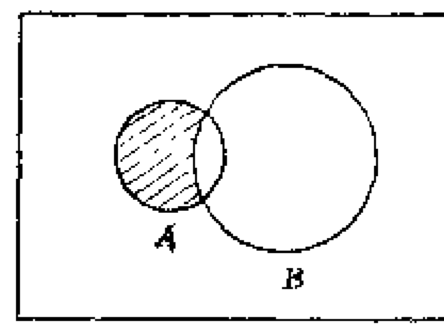


图 (1.2.6)

定义 1.2.6 事件A发生而事件B不发生，是一个事件，叫做事件A与事件B的差，记为 $A-B$，即

$$A-B=A\overline{B}$$

$A-B$ 是由属于A而又不属于B的基本事件所组成。

例 1.2.1 在检查某种圆柱形零件时，要求它的长度和直径都合格才算合格。我们从一批这种零件中随机取一个。设

$A=$“直径合格”

$B=$“长度合格”

$C=$“产品合格”

则

1. $C\subset A \quad C\subset B$
2. $\overline{C}=$“产品不合格”

   $\overline{B}=$“长度不合格”

   $\overline{A}=$“直径不合格”
3. $C=A\cap B$
4. $\overline{C}=\overline{A}\cup\overline{B}$
5. $C=A-\overline{B}$

例 1.2.2 在下面的电路图中，设

$A=$ “灯泡亮”

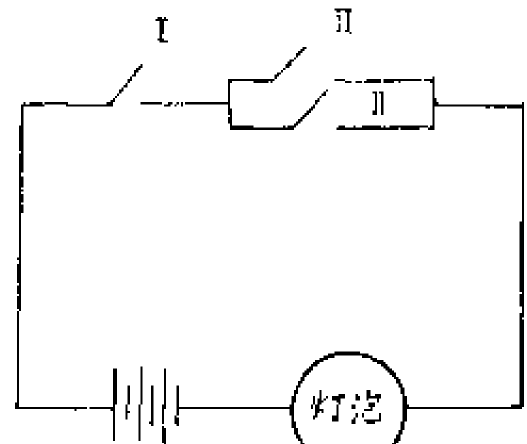
$B_1=$ “开关Ⅰ闭合”

$B_2=$ “开关Ⅱ闭合”

$B_3=$ “开关Ⅲ闭合”

则 $B_1B_2\subset A$，$B_1B_3\subset A$，

$B_1B_2\cup B_1B_3=A$。



图（1.2.7）

这是因为，如果 B_1B_2 发生，即开关Ⅰ，Ⅱ同时闭合，则整个电路接通，于是灯泡亮，即A发生．所以，$B_1B_2\subset A$。同理可说明为什么 $B_1B_3\subset A$。

为了说明 $B_1B_2\cup B_1B_3=A$，根据事件相等的定义，只要说明 $B_1B_2\cup B_1B_3\subset A$ 与 $A\subset B_1B_2\cup B_1B_3$ 同时成立即可。

如果 $B_1B_2\cup B_1B_3$ 发生，即 B_1B_2 或 B_1B_3 中至少一个发生，则整个电路接通，于是灯泡亮，所以

$$B_1B_2\cup B_1B_3\subset A \tag{1.2.1}$$

反之，如果A发生，即灯泡亮，那么 B_1B_2 或 B_1B_3 中至少有一个发生，所以

$$A\subset B_1B_2\cup B_1B_3 \tag{1.2.2}$$

由（1.2.1）和（1.2.2）得到 $B_1B_2\cup B_1B_3=A$。

§1.3 古典概率

一般地说，一个事件在每次试验中可能发生也可能不发生，这是具有偶然性的．但是，实践告诉我们，事件发生可能性的大小则是事件本身所具有的一种确定属性．不同的事件，发生的可能性的大小有所差别，有的发生的可能性大，有的发生的可能性小．比如，若一批产品中废品很多，则从中随机取一件得到废品的可能性较大，若一批产品中废品很少，则得到废品的可能性较小．了解事件发生的可能性的大小是有重要意义的．比如，为了掌握灯泡的生

产情况，就需要知道，在正常生产情况下，灯泡寿命在区间(1000—1500)中的可能性多大，寿命小于1500的可能性多大等等.

为了研究事件发生的可能性，就需要用一个数把这种可能性表达出来，给事件发生的可能性一个数值表示. 表示事件发生可能性大小的数值叫做事件的概率. 事件A，B，C，…的概率分别用$P(A)$，$P(B)$，$P(C)$，…表示.

那么，对于一个给定的事件A，概率$P(A)$到底是一个什么数呢？怎么来求呢？这一节我们先对一种最简单的情况加以说明. 先看下面两个例子.

1. 已知在1000件产品中，有3件废品，从这批产品中随机取出一件，“取得废品”的可能性多大？人们容易得出结论，这个可能性是3/1000.

2. 有10支电子管，编上了1至10的号码，现在从中随机取一个，很容易想到“取得偶数号码”的可能性是5/10，“取得的号码不大于2”的可能性是2/10，等等.

这两个例子具有下面两个共同特点：

(1) 试验的一切基本事件是有限个.

在第一个例子中设想1000件产品编上了1到1000的号码，所以随机取一个的一切基本事件有1000个. 在第二个例子中一切基本事件有10个.

(2) 一切基本事件是等可能的，即每一个基本事件发生的可能性都一样.

在第一个例子中的1000件产品，哪一件也不比其他件有什么特殊的被取出的可能性. 第二个例子中的10个电子管也同样是这样.

具有这两个特点的试验叫做古典型试验，称它的数学模型为古典概型. 由于它比较简单，所以是概率论初期研究的主要对象，概率的概念最初是对这种概型给出的. 定义如下.

定义1.3.1(古典概率定义)　对古典型试验，设试验的一切基本事件有n个，而事件A所包含的基本事件有k个，则A的概率定义为

$$P(A)=\frac{A\text{包含的基本事件数}}{\text{一切基本事件数}}=\frac{k}{n} \tag{1.3.1}$$

由于(1.3.1)只适用于古典型试验，所以由（1.3.1）定义的概率叫做古典概率。

在用这个定义求概率时，关键是要正确求出一切基本事件的总数n，其次是要正确求出事件A所包含的基本事件的个数k。下而举几个例子说明（1.3.1）的应用。

例 1.3.1　假设在10000双胶鞋中有40双次品，从其中随机取两双，问$A=$“两双中只有一双是次品”的概率是多少？

解　从10000双中取两双，所有可能的取法，即一切基本事件的总数是$\binom{10000}{2}$。一双合格品只能取之于10000双中的9960双，因此取一双合格品的取法有$\binom{9960}{1}$种，而取一双次品的取法有$\binom{40}{1}$种。所以两双中只有一双是次品的取法，即A所包含的基本事件的个数是9960×40。按（1.3.1)所求概率为

$$P(A)=(9960\times40)\Big/\binom{10000}{2}\doteq0.008$$

例 1.3.2　设有一批产品共N件，其中M件是废品，从N件中任意取出n件（$n<N$）。问$A=$“n件中有m件废品”的概率是多少？显然这是例1.3.1的一般化。

解　从N件中任取n件，所有可能取法有$\binom{N}{n}$种。m件废品的取法有$\binom{M}{m}$种，而$n-m$件好品的取法有$\binom{N-M}{n-m}$种，因此，在n件中有m件废品及$n-m$件好品的取法有$\binom{M}{m}\binom{N-M}{n-m}$种。则有

$$P(A)=\frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}}\qquad \begin{matrix}m\leqslant min\{M,n\}\\ n-m\leqslant N-M\end{matrix}$$

具有这种形式的概率计算，称其为服从超几何分布。它概括了比较广的一类问题。再看下例。

例 1.3.3 在20个人中有两名组长，现在把20个人随机分成两组，每组10人．问 A = “两组中各有一名组长”的概率是多少？

解 20个人随机分成各10人的两组，共有$\binom{20}{10}$种分法，在两名组长中取一名组长有$\binom{2}{1}=2$ 种取法，而在其余18人中取 9 名有$\binom{18}{9}$种取法，所以事件A包含基本事件数是 $2\binom{18}{9}$．因此

$$P(A)=2\binom{18}{9}\Big/\binom{20}{10}$$

例 1.3.4 将球一个接一个地放入n个盒子中，我们指定其中的一个盒，当这个指定的盒还空着时，我们就继续放，一旦，球放到了这个盒时，放球就停止．求在第k次球放到了这个盒子里的概率．

解 共放了k个球，由于每个球都有n种放法，所以k个球的放法数，即基本事件总数是n^k．第k次放到了指定的盒子中，而其余$k-1$次放到其余$n-1$个盒子中的放法数为 $(n-1)^{k-1}$，所以A= “恰好第k次把球放入指定的盒子中”的概率为

$$P(A)=(n-1)^{k-1}/n^k$$

有许多与例1.3.4相类似的问题，一般称其为“等待时间”问题．比如，一个人有n把钥匙，其中只有一把能打开某一扇门，由于他事先不知哪一把能打开门，所以他随机地用这些钥匙去试开，假定每把钥匙试完后又放回，问这个人恰好在第k次把门打开的概率是多少？结论与例1.3.4完全一样．所求概率为$(n-1)^{k-1}/n^k$．再比如，我们从人群中一个一个地随机选人，考察他们的生日，问恰好在第k次所选的人的生日为11月16日的概率是多少？答案为 $(364)^{k-1}/(365)^k$，这里365相当于例1.3.4中的n．

例 1.3.5 从0,1,2,…,9十个数码中随机而可重复地取出5个数码．求下列事件的概率：

A_1= “5个数码全相同”

A_2= “5个数码全不相同”

A_3 = "5 个数码中数码 3 出现两次"

解　由于取数码是可重复的，所以一切基本事件总数是10^5. A_1 中包含的基本事件数是10，所以

$$P(A_1)=10/10^5=1/10^4$$

A_2 中包含的基本事件数是 ${}_{10}P_5=10\times9\times8\times7\times6$，所以

$$P(A_2)=(10\times9\times8\times7\times6)/10^5$$

A_3 中包含的基本事件数是$\binom{5}{2}9^3$. 这是因为数码 3 在取出的 5 个数码中占有两个位置的方法有$\binom{5}{2}$种，而其余 3 个数码中的每一个都可从剩下的 9 个数码 0,1,2,4…,9 中取，在 9 个数码中可重复地取 3 个数码共有 9^3 种取法. 所以

$$P(A_3)=\binom{5}{2}9^3/10^5$$

例 1.3.6　有 r 个球，随机放在 n 个盒子中，($r\leqslant n$)，试求下列各事件的概率：

A_1 = "某指定的 r 个盒中各有一球"

A_2 = "恰有 r 个盒，其中各有一球"

A_3 = "某指定的一个盒中，恰有 k 个球"

解　（1） r 个球放入 n 个盒子里的方法共有 n^r 种，这就是基本事件的总数. 由于 r 个球在指定的 r 个盒中各放一个，共有$r!$ 种放法，所以A_1 包含的基本事件数为 $r!$ ，因此

$$P(A_1)=r!/n^r$$

（2） 由于在 n 个盒中选出 r 个盒的选法共有 $\binom{n}{r}$个，而对于每一种选法选出的 r 个盒，其中各放一个球的放法有 $r!$ 种，所以 A_2 包含的基本事件数为 $\binom{n}{r}r!$，因此

$$P(A_2)=\binom{n}{r}r!/n^r$$

（3） 由于在 r 个球中选出 k 个球的选法有 $\binom{r}{k}$种，而其余的$r-k$个球可任意地放在 $n-1$ 个盒子中，这种放法有$(n-1)^{r-k}$

种，所以 A_3 包含的基本事件数是$\binom{r}{k}(n-1)^{r-k}$，因此

$$P(A_3)=\binom{r}{k}(n-1)^{r-k}/n^r$$

与例1.3.6类似的问题还有许多，比如，把 r 个人随机地分配在 n 个房间里，求 A_1 = “某指定的 r 个房间中各有一人”，A_2 = “恰有 r 个房间其中各有一人”，A_3 = “某指定的房子中恰有 k 个人”的概率，结论与例1.3.6完全一样。

例1.3.6中的 A_2 与例1.3.5中的 A_2 实际上是不同形式的同一问题，即在可重复的排列中求不重复排列的概率。还可举出类似的问题。比如，设有 r 个人，$r\leqslant 365$。求 A = “r 个人的生日都不相同”的概率。可以认为 r 个人出生在365天中任何一天的可能性都相等。于是 r 个人相当于例1.3.6中的 r 个球，365天相当于例1.3.6中的 n 个盒子。所以所求概率为

$$P(A)=\frac{{}_{365}P_r}{365^r}=\frac{365!}{(365-r)!\,365^r}$$

例 1.3.7　一个袋中装有 N 个分别标上从1到 N 的球，从中有放回地取出 n 个。求 A = “取到的最大号码为 k”的概率。

解　一切基本事件总数为 N^n。所取最大号码为 k，意味着取出的号码要小于或等于 k，同时又不能都小于 k，即不能都小于或等于 $k-1$，所以 A 包含的基本事件数应为 $k^n-(k-1)^n$。因而所求概率为

$$P(A)=\frac{k^n-(k-1)^n}{N^n}$$

例 1.3.8　一个箱子里，有 n 双不同型号的鞋子。从中随机取出 $2r$ 只 $(2r<n)$，求下列事件的概率：

A_1 = “没有两只同型号的”

A_2 = “恰有一双同型号的”

解　从 n 双中随机取 $2r$ 只的方法有$\binom{2n}{2r}$种。

（1）若 $2r$ 只全是不同型号的，则它们必来自不同的 $2r$ 双。

在 n 双中取 $2r$ 双共有 $\binom{n}{2r}$ 种取法，对于所取的每一双可以取其左只，也可取右只，所以 A_1 包含的基本事件数是 $\binom{n}{2r}2^{2r}$，因此，

$$P(A_1)=\binom{n}{2r}2^{2r}\Big/\binom{2n}{2r}$$

（2）若 $2r$ 只中恰有一双是同型号的，则其中必有 $2r-2$ 只全是不同型号的，它们只能来自其余的 $n-1$ 双，在 $n-1$ 双中随机取 $2r-2$ 只全是不同型号的，由（1）知共有 $\binom{n-1}{2r-2}2^{2r-2}$ 种取法。又由于同型号的这一双可能是 n 双中的任一双，所以 A_2 所包含的基本事件数是 $n\binom{n-1}{2r-2}2^{2r-2}$，因此，

$$P(A_2)=n\binom{n-1}{2r-2}2^{2r-2}\Big/\binom{2n}{2r}$$

由（1.3.1）定义的古典概率具有下列基本性质。

1．对任意事件 A，有 $0\leqslant P(A)\leqslant 1$，

2．$P(U)=1$，

3．若事件 A 与事件 B 互斥，则有

$$P(A\cup B)=P(A)+P(B)$$

证明　由于任何事件 A 包含的基本事件数不能超过基本事件的总数，所以 1．成立；由于必然事件 U 包含一切基本事件，所以 2．成立。

现在证明 3．设 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_n\}$，$A=\{\omega_{11},\omega_{12},\cdots,\omega_{1k}\}$，$B=\{\omega_{21},\omega_{22},\cdots,\omega_{2l}\}$，则有 $P(A)=k/n$，$P(B)=l/n$。由于 A，B 互斥，所以 $\omega_{11},\omega_{12},\cdots,\omega_{1k}$ 与 $\omega_{21},\omega_{22},\cdots,\omega_{2l}$ 都不相同，因此 $A\cup B$ 包含 $k+l$ 个基本事件，按（1.3.1）则有 $P(A\cup B)=(k+l)/n=k/n+l/n$，故 3．成立。性质 3 叫做概率的可加性。

用归纳法不难把加法性质推广到任意有限个事件上去。即若 $A_1,A_2,\cdots,A_m$ 是 m 个两两互斥的事件，则有

$$P\left(\bigcup_{i=1}^{m}A_i\right)=P\left(\sum_{i=1}^{m}A_i\right)=\sum_{i=1}^{m}P(A_i)$$

由上面三个基本性质，还可导出下列性质：

4. $P(\overline{A})=1-P(A)$ (1.3.2)

证明 由于 $A\cup\overline{A}=U$，$A\cap\overline{A}=V$，所以

$$P(U)=P(A\cup\overline{A})=P(A)+P(\overline{A})=1$$

因此 (1.3.2) 成立。

(1.3.2) 对于计算概率是很方便的，有时直接计算概率 $P(A)$ 比较麻烦，而计算$P(\overline{A})$比较简单，这时可先求 $P(\overline{A})$，然后再用 (1.3.2) 求$P(A)$。

例 1.3.9 从0,1,2,…,9十个数码中随机而可重复地取出5个数码，求 A= "5个数码中至少有两个相同" 的概率。

解 事件A中包含的基本事件，情况比较复杂。它包括 "5个数码全相同"，"4个数码相同而与其余一个不同"，"3个数码相同而与其余两个不同"，等等。计算它们的个数比较麻烦。现在考虑事件$\overline{A}$= "5个数码全不相同"，由例1.3.5已知

$$P(\overline{A})=10\times9\times8\times7\times6/10^5=0.3024$$

所以根据 (1.3.2) 有

$$P(A)=1-P(\overline{A})=1-0.3024=0.6976$$

例1.3.10 r 个球随机放在 n 个盒子中，求 A= "至少有两个球放入同一盒中" 的概率。

解 A的逆事件为$\overline{A}$= "恰有 r 个盒各放有一个球"。由例1.3.6已知

$$P(\overline{A})=\binom{n}{r}r!/n^r$$

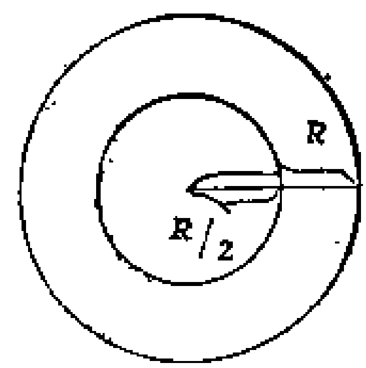
所以 $$P(A)=1-\binom{n}{r}r!/n^r$$

§1.4 几何概率

上述的概率古典定义只适用于古典概型，它要求试验的一切可能结果为有限且等可能。但是，在很多实际问题中，试验的一切可

能结果是无限个，这时古典概率定义就不适用了．考虑怎样把古典定义再进一步推广，把适用的范围再进一步扩大，比如适用于试验的一切可能结果是无限且等可能的情形，于是引出了几何概率定义．

先看一个例子．向半径为R的圆内随机掷一点，求事件A＝“点落在一个与该圆同圆心且半径为$R/2$的圆内”的概率．这时不论基本事件全体，还是A中的基本事件都是无限个，所以不能用公式（1.3.1）．在这里，假定所掷的点落在圆内任一点的可能性都一样．我们这样来理解这种等可能性的含义，就是点落在圆内某一区域g内的可能性与区域g的面积成正比，并且与g的位置无关．



图（1.4.1）

由于一切基本事件对应圆内所有点，而事件A包含的基本事件对应小圆内所有点，于是应有

$$\frac{P(A)}{P(\Omega)}=\frac{\text{小圆面积}}{\text{大圆面积}}$$

因此所求概率应当是

$$P(A)=\frac{\text{小圆面积}}{\text{大圆面积}}=\frac{1}{4}$$

将这个例子一般化，给出几何概率定义如下：

定义 1.4.1（几何概率定义）　设Ω是某一区域（可以是一维空间的，也可以是二维、三维空间的），向Ω中随机掷一点M，如果点M落在Ω中任一点是等可能的（或说是均匀分布），则说这个试验是几何型的．对于几何型试验，事件A＝“点M落在区域$g\subset\Omega$中”的概率定义为

$$P(A)=\frac{g\text{的测度}}{\Omega\text{的测度}} \tag{1.4.1}$$

这里测度是指长度、面积、体积等．

在实际中有很多试验，虽然并不是向某区域掷点，但是它的结

果可以用某区域Ω中的点表示，试验的一切结果对应区域Ω中的一切点，而且具有等可能性，这时（1.4.1）就可以用。下面举两个例子。

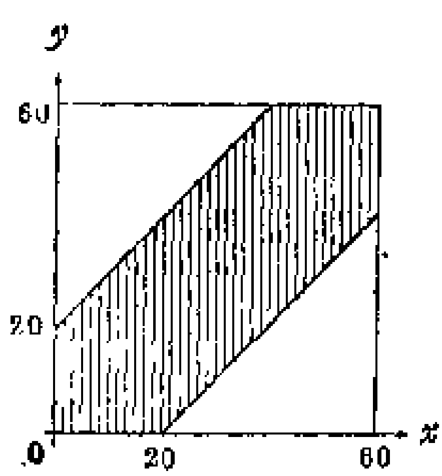
例 1.4.1（会面问题）两个人相约在7点到8点在某地会面，先到者要等候另一人20分钟，过时就离去。如果每人在指定的一小时内任一时刻到达是等可能的，求A=“两个人能会面”的概率。

解　以x表示甲到达时刻，y表示乙到达时刻，则两个人到达时间的一切可能结果对应边长为60的正方形里所有点。两个人能会面的充要条件是

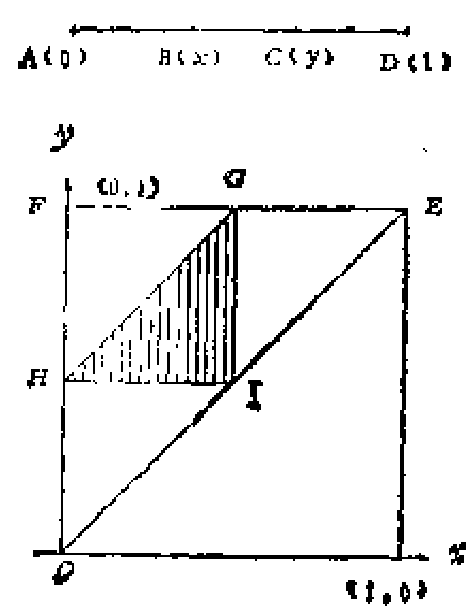
$$|x-y|\leqslant 20$$

所以事件A对应阴影部分里的一切点。根据（1.4.1），所求概率为

$$P(A)=(60^2-40^2)/60^2=5/9$$



图（1.4.2）



图（1.4.3）

例 1.4.2　在一线段AD上任取两点，将线段在此两点处折断，求A=“所得三折线能构成三角形”的概率。

解　设所取两点为B，C（B在C左边），并设线段长为1，点A，D坐标分别为0,1，点B的坐标为x，点C的坐标为y。则有

$$0\leqslant x\leqslant 1,\ 0\leqslant y\leqslant 1,\ x\leqslant y,$$

所以B，C两点的一切可能位置对应平面上三角形OEF里的所有

点，而 AB，BC，CD 构成三角形的充要条件是

$$|AB|+|BC|>|CD|$$
$$|BC|+|CD|>|AB|$$
$$|CD|+|AB|>|BC|$$

由于$|AB|=x$，$|AC|=y$，$|CD|=1-y$，代入上面不等式则得

$$y>\frac{1}{2} \quad x<\frac{1}{2} \quad y-x<\frac{1}{2}$$

所以使 AB，BC，CD 构成三角形的情形对应三角形 IGH 里的所有点。根据（1.4.1），所求概率为

$$P(A)=\frac{\triangle IGH\text{面积}}{\triangle OEF\text{面积}}=1/4$$

与古典概率类似，几何概率具有下列基本性质：

1．对任意事件 A，有 $0\leqslant P(A)\leqslant 1$；

2．$P(U)=1$；

3．若 $A_1,A_2,\cdots$两两互斥，则有

$$P\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i)$$

性质1，2由几何概率定义(1.4.1) 直接可得，性质3，叫做概率的可列可加性．它是根据长度、面积、体积等的可列可加性得到的．

§1.5 概率的公理化定义

（一）概率的统计定义

上面讲的古典概率及几何概率都有局限性，都要求试验的一切可能结果具有等可能性．可是，实际中有些试验是不具有等可能性的，比如，在一段时间内，一个电话交换台，可能“接到1次呼唤”、“接到2次呼唤”、“接到3次呼唤”，…．显然这些基本事件就不是等可能的，因此求“接到 i 次呼唤”的概率就不能用

(1.3.1) 和 (1.4.1)．这说明实际要求我们给出一个适用于一般试验的概率定义，于是在概率论的发展中，就进一步产生了概率的统计定义．这一定义的根据是：在试验的多次重复中，事件的频率具有稳定性．

首先说明什么是频率．假设在试验的n次重复中，事件A发生了k次，则k/n叫做事件A的频率．一个事件的频率不是一个固定的常数，这是因为在n次试验中该事件发生的次数k不是一个固定的常数，它随机地取 $0,1,2,\cdots$，n中的任何一个值．

但是，在试验的多次重复中，频率具有稳定性．为了说明频率的稳定性，看下面的例子．

在相似于大田的条件下，用同一品种的小麦种子作发芽试验，了解发芽的频率．我们分别抽取5粒，10粒，50粒，100粒，300粒，600粒种子进行试验，结果如下表：

抽取粒数	5	10	50	100	300	600
发芽粒数	5	8	44	91	272	542
发芽频率	1	0.8	0.88	0.91	0.907	0.903

可以看出，虽然抽取种子中的发芽数是不确定的，然而随着抽取粒数的增多，发芽的频率愈来愈明显地摆动在0.9左右．

还可以举出大量的试验资料来说明这种频率的稳定趋势．下面是我国某地1963年关于出生婴儿数的一个统计资料．

从这个材料可以看出，随着婴儿数的增加，女婴频率愈来愈明显地摆动在0.482左右．

尽管频率 k/n 本身是不确定的，但随着试验次数n的增加，频率愈来愈摆动在某一常数附近，n愈大，频率与这个常数出现大偏差的情况越稀少，呈现出一种稳定性，这种性质叫做频率稳定性．

频率能反映一个事件出现的频繁程度，从而能刻划事件出现的可能性的大小，而当试验次数很大时，频率与某一常数出现大偏差的可能性很小，因此，这个常数就可以作为事件的概率。概括说，概率的统计定义就是：

当试验次数n充分大时，事件的频率围绕某一常数摆动，这个常数就定义为该事件的概率。

月数	总数	男婴数	女婴数	女婴频率
第一个月	4418	2297	2121	0.4801
头两个月	10109	5251	4858	0.4806
头三个月	15183	7864	7319	0.4821
头四个月	19584	10112	9472	0.4837
头五个月	23745	12266	11479	0.4834
头六个月	28421	14667	13754	0.4839
头七个月	31269	16163	15106	0.4831
头八个月	35185	18160	16998	0.4831
头九个月	39537	20490	19047	0.4818
头十个月	44605	23078	21527	0.4826
头十一个月	49480	25615	23865	0.4823
十二个月	54434	28181	26253	0.4823

这个定义适用于一切类型的试验。根据这一定义，在实践中，我们可以把由大量观察所得到的频率作为概率的近似值。在很多情况下，这就足以满足实际需要了。

按上述概率的统计定义，概率实际上是频率的数学抽象，因此，频率所具有的性质，概率自然也应当有。下面就根据频率的性质推断出相应的概率的性质。

1. 对任意事件A，有$0\leqslant P(A)\leqslant 1$。

因为对任意事件A，频率必然总在0与1之间 $0\leqslant\frac{k}{n}\leqslant1$，因此$P(A)$也应在0与1之间，即$0\leqslant P(A)\leqslant1$。

2．$P(U)=1$

因为必然事件在每次试验中都一定发生，所以在n次试验中，它一定发生n次，因此频率$\frac{n}{n}=1$．所以应当有$P(U)=1$．

3．若事件A与事件B互斥，则有

$$P(A\cup B)=P(A)+P(B) \tag{1.5.1}$$

设在n次试验中A，B，$A\cup B$分别发生了k_A，k_B，$k_{A\cup B}$次，由于A与B互斥，所以$k_{A\cup B}=k_A+k_B$，因此有

$$\frac{k_{A\cup B}}{n}=\frac{k_A+k_B}{n}=\frac{k_A}{n}+\frac{k_B}{n}$$

与它对应的概率性质就应当是（1.5.1）．

用数学归纳法，可以把（3）推广到有限个事件上去，即若$A_1,A_2,\cdots,A_n$两两互斥，则有

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)$$

（二）概率的公理化定义

概率的统计定义，虽然它适用于一般情况，但它是不够精确的。在那里提到的"n充分大"，"围绕某个常数摆动"等都不是严格的数学概念，由统计定义所得到的概率基本性质并不是由定义证明得到的，而是根据频率性质推断出来的．因此，以这些作为概率理论的基础是不够的。为了使概率论建立在严格的理论基础之上，就需要对概率给出一个严格的一般的数学定义，因而产生了概率的公理化定义。

因为概率的概念与事件的概念有关，所以，为了给出一个确切的概率概念，应当把事件的概念再明确一下。

1．事件域

前面已经讲过，基本事件的全体构成基本事件空间Ω，事件是基本事件空间Ω的子集．在这里要强调一下，并非一定把Ω的一切子集都叫做事件，因为否则对讨论事件的概率会带来困难，同时我们所定义的事件又应当满足下面要求，即，若A是事件，则$\overline{A}$也是事件；若A，B是事件，则$A\cup B$也是事件；考虑到有时要研究无限个事件的运算，因此要求，若$A_1,A_2,\cdots$是事件，则$\bigcup_{i=1}^{\infty}A_i$也是事件．这样，我们所研究的事件是属于基本事件空间Ω的子集所构成的一个类$\mathscr{F}$，它满足下面三条．

（1）$\Omega\in\mathscr{F}$；

（2）若$A\in\mathscr{F}$，则$\overline{A}\in\mathscr{F}$；

（3）若$A_i\in\mathscr{F}$　$i=1,2,\cdots$，则$\bigcup_{i=1}^{\infty}A_i\in\mathscr{F}$

我们把$\mathscr{F}$叫做事件域．

定义 1.5.1（事件的定义）　若$\mathscr{F}$是由基本事件空间Ω的子集所构成的事件域，则$\mathscr{F}$中的元素叫做事件．

由上面三条可以证明$\mathscr{F}$还有以下性质．

（4）$\phi\in\mathscr{F}$，

（5）若$A_i\in\mathscr{F}$　$i=1,2,\cdots$，则$\bigcap_{i=1}^{\infty}A_i\in\mathscr{F}$，

（6）若$A_i\in\mathscr{F}$　$i=1,2,\cdots,n$，则$\bigcup_{i=1}^{n}A_i\in\mathscr{F}$，$\bigcap_{i=1}^{n}A_i\in\mathscr{F}$，

（7）若$A\in\mathscr{F}$　$B\in\mathscr{F}$，则$A-B\in\mathscr{F}$．

这四个性质，请读者自己证明．

2．概率的公理化定义

古典概率、几何概率及根据频率稳定性所定义的概率，具有共同的基本性质．我们就根据这些基本性质给出概率的公理化定义如下．

定义1.5.2（概率的定义）　设Ω为基本事件空间，$\mathscr{F}$为关于Ω的一个事件域，如果定义在事件域$\mathscr{F}$上的一个集函数$P(A)$ $(A\in\mathscr{F})$满足下面三条，则P叫做概率。

1．$P(A)\geqslant 0$　对一切$A\in\mathscr{F}$；

2．$P(\Omega)=1$，

3．若$A_i\in\mathscr{F}$ $i=1,2,\cdots$，互不相容，则

$$P\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i)$$

设Ω为基本事件空间，$\mathscr{F}$是关于Ω的事件域，P是概率，则三元总体$(\Omega,\mathscr{F},P)$叫做概率空间。概率空间是对一个试验的完整的数学描述。

利用概率的基本性质1.2.3，可以导出概率的其他性质。

性质1　$P(V)=0$

因为　$\Omega=\Omega\cup V\cup\cdots$所以由3．

$$P(\Omega)=P(\Omega)+P(V)+\cdots$$

因此　$P(V)=0$

性质2　概率具有有限可加性。如果$A_i\in\mathscr{F}$、$i=1,2,\cdots,n$，且$A_iA_j=V$，$i\neq j$，则

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)\tag{1.5.2}$$

在3．中取$A_{n+1}=A_{n+2}=\cdots=V$，利用$P(V)=0$，得

$$P\left(\bigcup_{i=1}^{n}A_i\right)=P\left(\bigcup_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i)=\sum_{i=1}^{n}P(A_i)$$

性质3　$P(\overline{A})=1-P(A)$

这在§1.3中已推导过。

性质4　若$A\subset B$，则$P(A)\leqslant P(B)$

因为　$B=A\cup(B-A)$，且$A\cap(B-A)=V$，则由1及(1.5.2)得

$$P(B)=P(A)+P(B-A)\geqslant P(A)$$

性质 5　$P(A)\leqslant 1$

因为对于任意事件A，恒有

$$A=AU\subset U$$

由性质 4　$P(A)\leqslant P(U)=1$.

性质 6　对任意两个事件A与B，有

$$P(A\cup B)=P(A)+P(B)-P(AB) \tag{1.5.3}$$

因为　$A\cup B=A\cup(B-A)$，且$(B-A)\cap A=V$，由(1.5.2) 得

$$P(A\cup B)=P(A)+P(B-A) \tag{1.5.4}$$

又因为　$B=(B-A)\cap AB$　且$(B-A)\cap AB=V$，再用 (1.5.2) 得

$$P(B)=P(B-A)+P(AB)$$

将此式代入 (1.5.4) ，即得 (1.5.3)

性质 7　$P(A\cup B)\leqslant P(A)+P(B)$

这是性质 6 的直接推论.

性质 8　如果 $A\supset B$，则

$$P(A-B)=P(A)-P(B) \tag{1.5.5}$$

因为当$A\supset B$时，$A=B\cup(A-B)$，且 $B\cap(A-B)=V$

所以　$P(A)=P(B)+P(A-B)$，移项即得 (1.5.5) .

性质 6 可推广到任意有限个事件上去，即，设 A_1,A_2，…，A_n为任意n个随机事件，则有

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)$$
$$+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)$$
$$-\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n). \tag{1.5.6}$$

这个公式叫做概率的一般加法公式，可用数学归纳法证明，见56页（2）.

概率还具有如下的连续性.

定理 1.5.1　设$\{A_n\}$为不减事件序列，即 $A_n\subset A_{n+1}$，$n=1$，$2,\cdots$，令　$A=\lim\limits_{n\to\infty}A_n=\bigcup\limits_{n=1}^{\infty}A_n$，则

$$P(A)=P(\lim_{n\to\infty}A_n)=\lim_{n\to\infty}P(A_n)$$

（即可将 lim 与 P 的顺序交换）称此性质为概率的下连续性。

证明　将 A 表示为互不相交的事件和

$$A=A_1+\left[\sum_{n=1}^{\infty}(A_{n+1}-A_n)\right]$$

由概率的可列可加性，有

$$P(A)=P(A_1)+\sum_{n=1}^{\infty}P(A_{n+1}-A_n)$$

即　$$P(A)=P(A_1)+\lim_{n\to\infty}\sum_{j=1}^{n-1}[P(A_{j+1})-P(A_j)]$$

因而有　$P(A)=\lim_{n\to\infty}P(A_n)$

定理 1.5.2　设$\{A_n\}$为不增事件序列，即 $A_n\supset A_{n+1}$，$n=1,2,\cdots$，令 $A=\lim_{n\to\infty}A_n=\bigcap_{n=1}^{\infty}A_n$，则

$$P(A)=P(\lim_{n\to\infty}A_n)=\lim_{n\to\infty}P(A_n)$$

称此性质为概率的上连续性。

证明　因为事件序列$\{\overline{A}_n\}$为不减序列，$\overline{A}=\bigcup_{n=1}^{\infty}\overline{A}_n$，由定理 1.5.1可得

$$P(\overline{A})=\lim_{n\to\infty}P(\overline{A}_n)$$

由此即可导出所求证的结果。

由性质 2 及定理 1.5.1 可知由概率的可列可加性可推得概率的有限可加性及连续性，反之，由有限可加性及连续性也能推得可列可加性，但只有有限可加性推不出可列可加性。

事实上，若 $A_i\in\mathscr{F}$，$i=1,2,\cdots$，互不相容，则由概率的有限可加性有

$$P\left(\sum_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)$$

等式左边对任意的n都不超过1，所以，右边的正项级数收敛，因此，应当有

$$\lim_{n\to\infty} P\left(\sum_{i=1}^{n} A_i\right)=\lim_{n\to\infty}\sum_{i=1}^{n} P(A_i)=\sum_{i=1}^{\infty}P(A_i) \quad (1.5.7)$$

若记 $S_n=\sum_{i=1}^{n}A_i$，则 $S_n\in\mathscr{F}$，$n=1,2,\cdots$，且 $S_n\subset S_{n+1}$，即$\{S_n\}$为$\mathscr{F}$中的一个单调不减事件序列，由概率的下连续性，有

$$\lim_{n\to\infty}P(S_n)=P(\lim_{n\to\infty}S_n)$$

即

$$\lim_{n\to\infty}P\left(\sum_{i=1}^{n}A_i\right)=P\left(\lim_{n\to\infty}\sum_{i=1}^{n}A_i\right)=P\left(\sum_{i=1}^{\infty}A_i\right) \quad (1.5.8)$$

由（1.5.7）及（1.5.8）自然有

$$P\left(\sum_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i) \quad (1.5.9)$$

如果只有有限可加性，而没有连续性，则只能得到（1.5.7），不可能得到（1.5.8），当然也不可能得到可列可加性（1.5.9）。

§1.6 条件概率与乘法公式

（一）条件概率与乘法公式

条件概率是一个重要的概念，为了给出这个概念，先看下面一个简单的例子。

假设有一批灯泡，共N个，其中有N_A个是合格品，有N_B个是甲厂生产的。在甲厂生产的N_B个灯泡中有N_{AB}个是合格品。从N个灯泡中随机取一个，设

$A=$“取得合格品”，$B=$“取得甲厂产品”

则 AB为“取得既是甲厂的又是合格品”

于是有

$$P(A)=\frac{N_A}{N},\ P(B)=\frac{N_B}{N},\ P(AB)=\frac{N_{AB}}{N}.$$

现在，如果已知取得的灯泡是甲厂生产的，问它是合格品的概率是多少？这概率相当于从甲厂生产的 N_B 个灯泡中随机取一个为合格品的概率，这个概率应为$\frac{N_{AB}}{N_B}$。一般它与 $P(A)$ 不同，因为试验的条件不同了，试验在原来条件下又增加了新的条件，即“已知取得产品是甲厂生产的。我们用一个新的符号 $P(A|B)$ 表示它，即

$$P(A|B)=\frac{N_{AB}}{N_B}.$$

它反映了“已知产品是甲厂的”这一条件限制，以区别没有这个条件限制时同一事件的概率 $P(A)$。读做“在 B 发生的条件下，A 发生的概率”，或简称为 A 对 B 的条件概率。$P(A|B)$ 与 $P(AB)$，$P(B)$ 显然有如下关系

$$P(A|B)=\frac{N_{AB}}{N_B}=\frac{\frac{N_{AB}}{N}}{\frac{N_B}{N}}=\frac{P(AB)}{P(B)} \tag{1.6.1}$$

上面我们对于一个古典型试验引出了 $P(A|B)$ 的表达式。对于一般的试验（1.6.1）是否也成立呢？由于概率是频率的抽象，所以，我们先看看频率是否有类似的关系。

设 A，B 为二事件，且 $P(B)>0$。将试验重复 n 次，若在 n 次中 A，B，AB 分别发生 k_A, k_B, k_{AB} 次，于是

$$B\text{的频率}=\frac{k_B}{n}$$

$$AB\text{的频率}=\frac{k_{AB}}{n}$$

因为在 B 发生的 k_B 次中，A 发生了 k_{AB} 次，所以

$$A\text{对}B\text{的频率}=\frac{k_{AB}}{k_B}$$

由于 $P(B)>0$，所以当 n 充分大时$\frac{k_B}{n}>0$，所以有

$$\frac{k_{AB}}{k_B} = \frac{\frac{k_{AB}}{n}}{\frac{k_B}{n}}$$

就是说，频率具有与（1.6.1）类似的性质. 所以在一般情况下，概率也应具有性质（1.6.1）. 因此给出条件概率定义如下.

定义 1.6.1（条件概率定义） 设 B 为任意事件，$P(B)>0$，则在 B 发生的条件下，事件 A 的条件概率 $P(A|B)$ 定义为

$$P(A|B)=\frac{P(AB)}{P(B)} \tag{1.6.2}$$

由（1.6.2）立刻得到

$$P(AB)=P(B)P(A|B) \tag{1.6.3}$$

这个等式叫做概率的乘法公式. 这个公式很容易推广到 n 个事件上去.

设 $A_1,A_2,\cdots,A_n$ 为 n 个事件，$n\geqslant 2$，满足 $P(A_1A_2\cdots A_{n-1})>0$，则有

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2|A_1)P(A_3|A_1A_2)\cdots P(A_n|A_1A_2\cdots A_{n-1}) \tag{1.6.4}$$

证明 由于 $P(A_1)\geqslant P(A_1A_2)\geqslant\cdots\geqslant P(A_1A_2\cdots A_{n-1})>0$，所以

$$P(A_1A_2\cdots A_n)=P(A_1)\cdot\frac{P(A_1A_2)}{P(A_1)}\cdot\frac{P(A_1A_2A_3)}{P(A_1A_2)}\cdots\frac{P(A_1A_2\cdots A_n)}{P(A_1A_2\cdots A_{n-1})}$$

对上式各项用（1.6.2），便得（1.6.4）

例 1.6.1 已知某工厂生产的产品的合格率为0.96，而合格品中的一级品率为0.75，求该厂生产的产品是一级品的概率.

解 设 $A=$ “产品是一级品”

$B=$ “产品是合格品”

因为如果产品是一级品则必是合格品，所以在本例中 $P(A)=P(AB)$，由题意已知

$$P(B)=0.96, \qquad P(A|B)=0.75$$

因此根据（1.6.3）有

$$P(A)=P(AB)=P(B)P(A|B)=0.96\times 0.75=0.72$$

例 1.6.2 袋中放有5个黑球和4个白球，从中随机取出一个，然后放回，并同时再加进与抽出的球同色的球2个，再取第二个．问所取两球都是白球的概率．

解 设 $A=$"第一次取得白球"

$B=$"第二次取得白球"

则 $AB=$"所取两个都是白球"

由于 $P(A)=4/9 \qquad P(B|A)=6/11$

所以 $P(AB)=P(A)P(B|A)=4/9\times 6/11=8/33$

例 1.6.3 袋中放有 a 个白球和 b 个黑球．随机取出一个，然后放回，并同时再放进与取出的球同色的球 c 个，再取第二个，这样，连续取3次．问取出的3个球中头两个是黑球、第3个是白球的概率是多少？

解 设 $A_1=$"第一次取得黑球"

$A_2=$"第二次取得黑球"

$A_3=$"第三次取得白球"

由于 $P(A_1)=b/(a+b) \qquad P(A_2|A_1)=(b+c)/(a+b+c)$

$P(A_3|A_1A_2)=a/(a+b+2c)$

所以所求概率为

$$P(A_1A_2A_3)=P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|A_1A_2)$$
$$=b/(a+b)\times(b+c)/(a+b+c)\times a/(a+b+2c)$$

条件概率的概念是相对的，它是对某确定的试验来说的．对某个试验来说的条件概率，当试验的条件改变时，它又可能成为无条件概率．比如在本节一开始讲的灯泡的例子中，$\frac{N_{AB}}{N_B}$是在已知所取的灯泡是甲厂的条件下，取得合格品的概率．若我们的试验是从整批灯泡中随机取一个，则$\frac{N_{AB}}{N_B}$就是"取得合格品"的条件概率．若

我们一开始就只考虑甲厂生产的 N_B 个灯泡，我们的试验是从 N_B 个灯泡中随机取一个，则$\frac{N_{AB}}{N_B}$就是“取得合格品”的无条件概率.

定理 1.6.1 如果 $P(B)>0$，则条件概率$P(A|B)(A\in\mathscr{F})$具有以下三个基本性质.

（1）对于每一个$A\in\mathscr{F}$，$P(A|B)\geqslant 0$

（2）$P(\Omega|B)=1$

（3）若 $A_i\in\mathscr{F}$，$i=1,2,\cdots$ 互不相容，则有

$$P\left(\bigcup_{i=1}^{\infty} A_i \middle| B\right)=\sum_{i=1}^{\infty} P(A_i|B)$$

证明（1）因为 $P(AB)\geqslant 0$，$P(B)>0$,所以，由（1.6.2）可知 $P(A|B)\geqslant 0$.

（2）$$P(\Omega|B)=\frac{P(\Omega B)}{P(B)}=\frac{P(B)}{P(B)}=1$$

（3）$$P\left(\bigcup_{i=1}^{\infty} A_i \middle| B\right)=\frac{P\left(\bigcup_{i=1}^{\infty} A_i B\right)}{P(B)}$$

$$=\sum_{i=1}^{\infty}\frac{P(A_i B)}{P(B)}=\sum_{i=1}^{\infty} P(A_i|B)$$

由定理 1.6.1 可知,$P(A|B)(A\in\mathscr{F})$作为 A 的集函数是$\mathscr{F}$上的概率. 既然条件概率也满足概率定义的三个条件，所以，对概率所证明的结果对于条件概率也适用.

（二）全概率公式

在概率的计算中，我们总是希望从已知事件的概率去求所考虑的事件的概率，或从比较简单的事件的概率，去求未知的比较复杂的事件的概率. 在这方面，下面讲的全概率公式和贝叶斯公式起着很重要的作用.

在给出全概率公式之前，先分析下面的例子.

例 1.6.4 某工厂有甲、乙、丙三台机器，它们的产量分别占总产量的0.25，0.35，0.40，而它们的产品中的废品率分别为0.05，0.04，0.02。今从总产品中随机取一件，求所取产品为废品的概率。

解 设 B= "取出的产品为废品"

如果我们知道产品是来自哪一台，则产品为废品的概率由题意便知。问题是不知道它来自哪一台，只知道它必来自三台中的一台，因此来自三台的可能性都要考虑。设

A_1= "所取产品来自甲台"

A_2= "所取产品来自乙台"

A_3= "所取产品来自丙台"

则 $B=A_1B\cup A_2B\cup A_3B$，由于 A_1,A_2,A_3 互不相容，所以

$$P(B)=P(A_1B)+P(A_2B)+P(A_3B)$$

而

$$P(A_1B)=P(A_1)P(B|A_1)$$
$$P(A_2B)=P(A_2)P(B|A_2)$$
$$P(A_3B)=P(A_3)P(B|A_3)$$

已知

$$P(A_1)=0.25 \quad P(A_2)=0.35 \quad P(A_3)=0.40$$
$$P(B|A_1)=0.05 \quad P(B|A_2)=0.04 \quad P(B|A_3)=0.02$$

所以，所求概率为

$$P(B)=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)+P(A_3)P(B|A_3)$$
$$=0.0345$$

将此题解法一般化，便得到全概率公式。

定理 1.6.2 设 $A_1,A_2,\cdots$ 是有限个或可列个互不相容的事件，而且它们的和是整个基本事件空间，$\bigcup_{i=1}^{\infty}A_i=U$，$P(A_i)>0$，$i=1,2,\cdots$，则对任一事件 B，有

$$P(B)=\sum_{i=1}^{\infty}P(A_i)P(B|A_i) \tag{1.6.5}$$

(1.6.5) 式叫做全概率公式.

证明　由假设可知事件 B 必与某一 A_i 同时发生，即

$$B = A_1B \cup A_2B \cup \cdots$$

由于 A_i 两两互斥，所以 A_iB 也两两互斥，由概率的可列可加性，得

$$P(B) = \sum_{i=1}^{\infty} P(A_iB)$$

再利用 (1.6.3)，则得

$$P(B) = \sum_{i=1}^{\infty} P(A_i)P(B|A_i)$$

当　$P(A_i)$，$P(B|A_i)$ 已知时，或它们比较容易求时，则可用 (1.6.5) 求 $P(B)$.

例 1.6.5　设门被锁的概率是$\frac{1}{2}$，在12把钥匙中有一把能打开门. 如果某人从12把钥匙中随取出 2 把去开门，问此人能打开门的概率是多少？

解　设 B = "所取 2 把能打开门"

A_1 = "门被锁"

A_2 = "门未锁"

显然 A_1, A_2, B 符合全概率公式条件. 由于

$$P(A_1) = P(A_2) = \frac{1}{2} \quad P(B|A_1) = \binom{11}{1} \Big/ \binom{12}{2} = 1/6$$

$$P(B|A_2) = 1,$$

所以由 (1.6.5) 有

$$P(B) = \frac{1}{2} \times \frac{1}{6} + \frac{1}{2} = 7/12$$

(三) 贝叶斯公式

先看下面例子.

例 1.6.6　问题的条件与例 1.6.4 同，若从总产品中随机取一

件，结果为废品，求这个废品是甲台机器生产的概率。

解 仍用例1.6.4中的符号，则所求概率为条件概率 $P(A_1|B)$，由条件概率定义可知

$$P(A_1|B)=\frac{P(A_1B)}{P(B)}=\frac{P(A_1)P(B|A_1)}{\sum_{i=1}^{3}P(A_i)\,P(B|A_i)}$$

上面等式右边分子分母中所出现的概率，由题意都已知，代入便得

$$P(A_1|B)=\frac{0.25\times 0.05}{0.0345}\doteq 0.362$$

将此题解法一般化，便得到很有用的贝叶斯（*Bayes*）公式。

定理 1.6.3 设 $A_1,A_2,\cdots$，是有限个或可列个互不相容的事件，且 $\bigcup_i A_i=U$，$P(A_i)>0$，$i=1,2,\cdots$，则对任一事件 B，$P(B)>0$，有

$$P(A_i|B)=\frac{P(A_i)P(B|A_i)}{\sum_{i=1}^{\infty}P(A_i)P(B|A_i)} \tag{1.6.6}$$

公式（1.6.6）叫做贝叶斯公式，有时也叫后验概率公式。

证明 根据条件概率定义有

$$P(A_i|B)=\frac{P(A_iB)}{P(B)}$$

对分子分母分别用乘法公式及全概率公式即得（1.6.6）

这个公式只不过是把条件概率表达式 $P(A_i|B)=\frac{P(A_iB)}{P(B)}$ 换了一个写法，这样理解便于记忆。当然这个公式的作用是很重要的。它常用于解决下面一类问题：如果某事件 B 能而且只能与两两互斥的 n 个事件 $A_1,A_2,\cdots,A_n$ 之一同时发生，而且由于某种理由，我们在试验之前知道了 $P(A_i)$ 及 $P(B|A_i)$，$i=1,2,\cdots,n$，现在做了一次试验，假定事件 B 发生了，这时如何根据 B 发生这一新的条件

来求事件A_i发生的概率$P(A_i|B)$. 它的作用就在于在我们求A_i发生的可能性时充分利用了B已发生这一情况所带来的信息. 与后验概率$P(A_i|B)$相对，有时把$P(A_i)$叫做先验概率.

例 1.6.7 发报机发出信号0或1，由于随机干扰，发出信号0（或1）并不一定收到0（或1）. 假定已知发出0而收到正好是0的概率为0.95，发出1而收到正好是1的概率为0.90，又已知发出信号0的概率是0.4(如右图). 如果已知收到信号是1，求发出信号是1的概率.

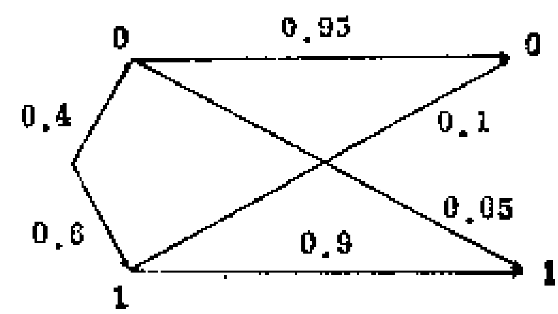


图 (1.6.1)

解 设 A_1= "发出信号1"

A_2= "发出信号0"

B= "收到信号1"

已知 $P(A_1)=0.6$ $P(A_2)=0.4$ $P(B|A_1)=0.90$

$P(B|A_2)=0.05$，所求概率为

$$P(A_1|B)=\frac{P(A_1)P(B|A_1)}{P(A_1)P(B|A_1)+P(A_2)P(B|A_2)}$$

$$=\frac{0.6\times0.9}{0.6\times0.9+0.4\times0.05}=\frac{27}{28}$$

§1.7 事件的独立性

(一) 两个事件的独立性

直观地说，若两个事件A与B，其中任何一个发生与否都不影响另一个发生的可能性，我们就说A与B相互独立.

比如，两个人各对靶射击一次，用A表示事件"甲击中"，用B表示事件"乙击中"，显然A与B相互独立.因为"乙击中"的可

能性，并不因为甲是否击中而有改变，同样“甲击中”的可能性也不因为乙是否击中而有改变．

上面关于两个事件相互独立的概念是描述性的，为了理论上的需要，我们给出两事件相互独立的数学定义．

因为若 A 与 B 独立，显然应当有

$$P(A|B)=P(A)\qquad P(B|A)=P(B)$$

由乘法公式，则应当有

$$P(AB)=P(A)P(B)$$

定义 1.7.1 对事件 A 与 B，若有

$$P(AB)=P(A)P(B) \tag{1.7.1}$$

则称 A 与 B 相互独立．

由定义 1.7.1 可知关于独立性有下面两个性质．

性质 1 如果 $P(B)>0$，则（1.7.1）等价于

$$P(A)=P(A|B)$$

证明 若 $P(A)=P(A|B)$，则由乘法公式 $P(AB)=P(B)P(A|B)$ 可得 $P(AB)=P(A)P(B)$

反之，若 $P(AB)=P(A)P(B)$，则

$$P(A)P(B)=P(AB)=P(B)P(A|B)$$

由于 $P(B)>0$，两边同除以 $P(B)$，即得 $P(A)=P(A|B)$

虽然 $P(B)>0$ 时，$P(A)=P(A|B)$ 与（1.7.1）等价，但仍用（1.7.1）作为两个事件独立的定义．这是因为它具有对称性，反映了独立性的相互性质，并且（1.7.1）对 $P(A)=0$，$P(B)=0$ 时也适用．

性质 2 若 A 与 B 相互独立，则 A 与 $\overline{B}$，$\overline{A}$ 与 B，$\overline{A}$ 与 $\overline{B}$ 也相互独立．

证明 已知 $P(AB)=P(A)P(B)$，所以

$$\begin{aligned}P(A\overline{B})&=P(A-AB)=P(A)-P(AB)\\&=P(A)-P(A)P(B)\\&=P(A)(1-P(B))=P(A)P(\overline{B})\end{aligned}$$

即A与$\overline{B}$相互独立。既然由A与B相互独立推出了A与$\overline{B}$相互独立，同样由A与$\overline{B}$相互独立也可推出$\overline{A}$与$\overline{B}$相互独立。又由于$B=\overline{\overline{B}}$，所以由$\overline{A}$与$\overline{B}$相互独立，又可推出$\overline{A}$与$B$相互独立。

（二）n个事件的独立性

关于独立性的概念可以推广到n个事件上去。

定义 1.7.2　设$A_1,A_2,\cdots A_n$是n个事件，如果对n个事件中的任意k个事件$A_{i_1}, A_{i_2}, \cdots A_{i_k}(1\leqslant i_1<i_2<\cdots<i_k\leqslant n)$有

$$P(A_{i_1}A_{i_2}\cdots A_{i_k})=P(A_{i_1})P(A_{i_2})\cdots P(A_{i_k}) \quad (1.7.2)$$

则称$A_1,A_2,\cdots,A_n$相互独立。

(1.7.2) 式实际上代表了下面这些式子。

$$P(A_{i_1}A_{i_2})=P(A_{i_1})P(A_{i_2}) \quad 共\binom{n}{2}个$$

$$P(A_{i_1}A_{i_2}A_{i_3})=P(A_{i_1})P(A_{i_2})P(A_{i_3}) \quad 共\binom{n}{3}个$$

$$\cdots\cdots$$

$$P(A_1A_2\cdots A_n)=P(A_1)P(A_2)\cdots P(A_n) \quad \binom{n}{n}=1个$$

总共有$\sum_{i=2}^{n}\binom{n}{i}=(1+1)^n-\binom{n}{1}-\binom{n}{0}=2^n-n-1$个式子。

由定义1.7.2可知，如果$A_1,A_2,\cdots,A_n$相互独立，则其中任意m个$(m<n)$事件也相互独立。

在n个事件相互独立的概念中，应特别注意，由$A_1,A_2,\cdots,A_n$中两两事件的独立性，不能推出它们的相互独立性。看下面例子。

例 1.7.1　设基本事件空间为$\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4\}$，$P(\{\omega_i\})=\frac{1}{4}$，$i=1,2,3,4$，又设　$A=\{\omega_1,\omega_2\}$，$B=\{\omega_1,\omega_3\}$。$C=\{\omega_1,\omega_4\}$。则有

$$P(A)=P(B)=P(C)=\frac{1}{2}$$

$$P(AB)=P(AC)=P(BC)=P(\{\omega_1\})=\frac{1}{4}$$

$$P(ABC)=P(\{\omega_1\})=\frac{1}{4}$$

可以看出 $P(AB)=P(A)\cdot P(B)$

$$P(AC)=P(A)\cdot P(C)$$

$$P(BC)=P(B)\cdot P(C)$$

$$P(ABC)\neq P(A)\cdot P(B)\cdot P(C)$$

就是说A，B，C两两独立，但是总起来不相互独立。

例 1.7.2 设 $\Omega=\{\omega_1,\omega_2,\cdots,\omega_8\}$，$P(\{\omega_i\})=\frac{1}{8}$，$i=1,2,\cdots,8$，又设 $A=\{\omega_1,\omega_2,\omega_3,\omega_4\}$，$B=\{\omega_1,\omega_2,\omega_3,\omega_5\}$，$C=\{\omega_1,\omega_6,\omega_7,\omega_8\}$，则有

$$P(A)=P(B)=P(C)=\frac{1}{2}$$

$$P(AB)=\frac{3}{8},\quad P(AC)=\frac{1}{8},\quad P(BC)=\frac{1}{8},\quad P(ABC)=\frac{1}{8}$$

可以看出

$$P(AB)\neq P(A)\cdot P(B)$$

$$P(AC)\neq P(A)\cdot P(C)$$

$$P(BC)\neq P(B)\cdot P(C)$$

$$P(ABC)=P(A)\cdot P(B)\cdot P(C)$$

这两个例子说明，在n个事件相互独立定义中的2^n-n-1个式子缺一不可。

应用（1.7.2）可以得到一个计算相互独立事件和的概率的简便方法。

设 $A_1,A_2,\cdots,A_n$ 是n个相互独立的事件，则有

$$P(A_1\cup A_2\cup\cdots\cup A_n)=1-P(\overline{A_1})P(\overline{A_2})\cdots P(\overline{A_n})\quad(1.7.3)$$

这是因为

$$\overline{A_1\cup A_2\cup\cdots\cup A_n}=\overline{A_1}\cap\overline{A_2}\cap\cdots\cap\overline{A_n}$$

所以

$$P(A_1\cup A_2\cup\cdots\cup A_n)=1-P(\overline{A_1}\cap\overline{A_2}\cap\cdots\cap\overline{A_n})$$

又由于$\overline{A_1}$，$\overline{A_2},\cdots,\overline{A_n}$相互独立，所以得（1.7.3）

例 1.7.3 一个人看管三台机床，设在任一时刻机床正常工作

（不需要人照顾）的概率对第一台是0.9，对第二台是0.8，对第三台是0.85，求在任一时刻，（1）三台机床都正常工作的概率；（2）三台机床中至少有一台正常工作的概率.

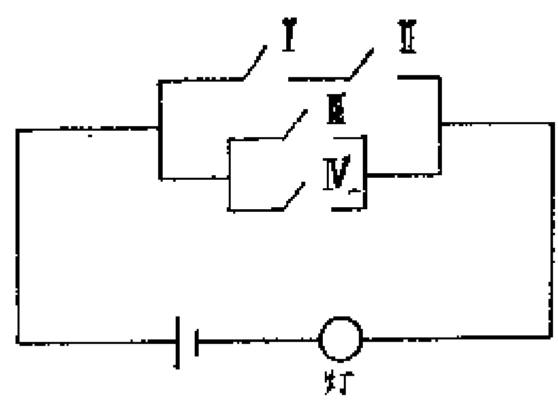
解　我们可以认为三台机床的工作正常与否是相互独立的，用 A_i 表示事件"第 i 台机床正常工作"．则

（1）所求概率为

$$P(A_1A_2A_3)=P(A_1)P(A_2)P(A_3)$$
$$=0.9\times0.8\times0.85=0.612$$

（2）所求概率为

$$P(A_1\cup A_2\cup A_3)$$
$$=1-P(\overline{A_1})P(\overline{A_2})P(\overline{A_3})$$
$$=1-(0.1\times0.2\times0.15)$$
$$=0.997$$



图（1.7.1）

例 1.7.4　在右图所示的开关电路中，开关Ⅰ，Ⅱ，Ⅲ，Ⅳ的开（或关）的概率均独立地等于 $\frac{1}{2}$，求灯亮的概率.

解　设 A_1,A_2,A_3,A_4 分别表示开关Ⅰ，Ⅱ，Ⅲ，Ⅳ关闭，则"灯亮" $=A_1A_2\cup A_3\cup A_4$．故所求概率为

$$P(A_1A_2\cup A_3\cup A_4)=1-P(\overline{A_1A_2})P(\overline{A_3})\,P(\overline{A_4})$$
$$=1-\{[1-P(A_1A_2)]\,P(\overline{A_3})\,P(\overline{A_4})\}$$
$$=1-\left[1-\left(\frac{1}{2}\right)^2\right]\cdot\frac{1}{2}\cdot\frac{1}{2}$$
$$=1-3/16=13/16$$

习　　题

1　试求对应下述试验的基本事件空间.

（1）同时扔3枚可区别的硬币；

（2）袋中放有4个可区别的白球和6个可区别的红球，从中随机取出一个。

2 设$\Omega=\{\omega_1,\omega_2,\cdots,\omega_{12}\}$，$A=\{\omega_1,\omega_2,\cdots,\omega_6\}$，$B=\{\omega_4,\omega_5,\cdots,\omega_9\}$，试求下列事件所包含的基本事件。

（1）$A\cap\overline{B}$，（2）$(A\cap\overline{B})\cup(\overline{A}\cap B)$，（3）$\overline{A\cup B}$，

（4）$\overline{A}\cap\overline{B}$，（5）$\overline{A\cap B}$，（6）$\overline{A}\cup\overline{B}$。

3 随机抽检3件产品，设

$A=$"三件中至少有一件是废品"

$B=$"三件中至少有两件是废品"

$C=$"三件都是好品"

问$\overline{A}$，$\overline{B}$，$A\cup C$，$A\cap C$，$A-B$各表示什么事件？

4 用A，B，C的式子将下列诸事件表示出来。在A，B，C三个事件中

（1）只有A发生；（2）只有A不发生；

（3）三个事件都发生；（4）三个事件都不发生；

（5）只有一个发生；（6）只有两个发生；

（7）至少有一个发生；（8）至少有两个发生；

（9）至多有两个发生。

5 证明下列等式：

（1）$\overline{A\cup B}=\overline{A}\cap\overline{B}$ （2）$\overline{A\cap B}=\overline{A}\cup\overline{B}$

（3）$(A\cup B)-B=A\cap\overline{B}$ （4）$(A\cup B)-AB=A\overline{B}\cup\overline{A}B$

（5）$A\cup A=A$，$A\cap A=A$ $\overline{\overline{A}}=A$

6 在100粒大豆中，有10粒虫豆，90粒好豆。从这100粒大豆中随机取出5粒，求"其中恰有3粒虫豆"的概率。

7 有50件产品，已知其中有4件不合格。从中随机取出3件，求"3件中至少有一件是不合格品"的概率。

8 袋中放有4个红球，3个白球和5个黑球，从中随机取出5个球，假定取球是有放回的（即每取出一球后将其放回，然后再取下一个）。求"取得两个红球和3个白球"的概率。

9　全班共有36人，其中男生、女生各一半．现在将其随机分成两组，每组各18人．求“两组中男生、女生各一半”的概率．

10　袋中有10个纸片，上面分别记上1到10的十个数码，随机从中取出3个．求

（1）“三个中最小数码是5”的概率；

（2）“三个中最大数码是5”的概率．

11　随机而可重复地由0,1,2,…9十个数码中取出 k 个数码（有顺序）．求“k 个数码中既无0又无1”的概率．

12　教室中有 $n\leqslant 12$ 名学生．求“至少有两名学生的生日在同一个月”的概率．

13　在1,2,3,4,5五个数码中不放回地随机取两个数码（有顺序）．求下列事件的概率：

A_1＝“第一次取得奇数”

A_2＝“第二次取得奇数”

A_3＝“两次都取得奇数”

14　将数码1,2,3随机排成一列，求“至少有一个数码与它所占位置的号数一致”的概率．

15　袋中放有分别记1到10的十个小纸片，不放回地从中随机取出两个，求“两纸片上的数的和等于10”的概率：

16　从数码1,2,…,50中随机取出一个，求“能被6或8整除”的概率．

17　从自然数1,2,…，n 中随机取出两个，求“两个数是相连的数”的概率．

（1）不放回地取　　　　（2）有放回地取

18　一人有 n 把钥匙，其中只有一把能打开门．他逐个试验（注意与例1.3.4不一样，这里的抽取是不放回的），求“恰在第 k 次试验时打开门”的概率．

19　将数1,2,…,n 依任意顺序排列，求

A_1＝“在数列中1与2相继出现”

A_2 = "在数列中 1， 2 及 3 相继出现" 的概率.

20 从 1,2,3,4,5 五个数码中，随机而可重复地取25个数码，求这25个数码中各数码都取到 5 次的概率.

21 r 个球放入 n 个盒子中，求"某指定的盒子是空的" 的概率.

22 求 n 个人的生日互不相同的概率($n \leqslant 365$).

23 袋中放有 a 个黑球 b 个白球，从中不放回地一个一个随机往外取，求"第 k 次取得黑球" 的概率.

24 袋中放有 a 个黑球 b 个白球，从中有放回地一个一个随机地往外取，直到出现 k 个白球为止，求"必须抽取的次数为 r" 的概率.

25 长为 L 的电话线 OB，在某一点 C 被切断，求"点 C 与点 O 的距离不小于 l" 的概率.

26 在时间间隔 T 内的任何瞬间，两信号都是等可能地进入收音机. 如果两个信号的时间间隔小于 τ，则收音机受到干扰. 求"收音机受到干扰" 的概率.

27 在长为 l 的线段 OD 上随机取三点 A, B, C，求"三线段 OA,OB,OC能构成三角形" 的概率.

28 甲、乙两艘轮船均欲靠同一码头，而两艘轮船到达时间互相独立，并且在一昼夜内是等可能的. 如果轮船甲需要停留时间为 1 小时，而轮船乙停留时间为 2 小时. 求"一艘轮船需要等待空出码头" 的概率.

29 两人约定 7 点到 8 点在某地会面，求"一人要等另一人半小时以上" 的概率.

30 在半径为 R 的圆内画平行线，如果这些弦与垂直于弦的直径的交点在该直径上的位置是等可能的. 求"任意画的弦的长度大于 R" 的概率.

31 袋中有 a 个白球 b 个黑球，从中不放回地一个一个地随机取出 3 个，求"3 个都是白球" 的概率.

32　袋中放有 a 个白球 b 个黑球，随机取出一个，然后放回，并同时加进 C 个与所取的球同样颜色的球和 d 个相反颜色的球（这时袋中有 $a+b+c+d$ 个球了），再取下一个，这样连续取 3 次。求"3 个球中头两个是黑球第 3 个是白球"的概率。

33　已知 $P(A)=0.3$，$P(B)=0.4$，$P(A|B)=0.32$，求下面的概率：

$$P(A\cup B),P(A\cap B)\ P(\overline{A\cup B}),\ P(\overline{A\cap B})$$

34　有两个袋子，一个里放有 5 个白球 7 个红球，另一个里放有 4 个白球 2 个红球。现在从两个袋子中随机取一个袋子，然后从所取的袋子中随机取一个球，求"取得白球"的概率。

35　袋 1 中放有 x 个白球 y 个红球，袋 2 中放有 u 个白球 v 个红球。从袋 1 中随机取一个球放入袋 2 中，然后从袋 2 中随机取一个球。求"所取球为白球"的概率。

36　两个坏零件和两个好零件放在一起，从中随机逐个往外取（不放回），直到两个坏的均取出为止，求下列事件的概率：

（1）"共取了 2 次"　　（2）"共取了 3 次"

（3）"共取了 4 次"

37　有 5 个坛子，其中有

2 个坛子里有 2 个白球 1 个黑球；

1 个坛子里有10个黑球；

2 个坛子里有 3 个白球 1 个黑球。

从 5 个坛子中随机取出一个，然后从其中随机取出一个球，求"取得白球"的概率。

38　设 $A_i(i=1,2,\cdots,n)$，B，C是具有正概率的事件，$\bigcup_{i=1}^{n} A_i=U$，A_i互斥，且$P(A_i)$，$P(B|A_i)$，$P(C|B\cap A_i)$已知，求概率 $P(C|B)$。

39　发报机发出信号 0 或 1，设

$A=$"发出信号 1"

$B=$"收到信号 1"

已知$P(A)=0.5, P(B|A)=P(\overline{B}|\overline{A})=0.90$.求$P(A|B), P(A|\overline{B})$.

40 设有5个袋子里面放有黑、白球，其中1号袋中白球占1/3，另外2,3,4,5号4个袋子中白球都占1/4。今从中随机取一个袋子，从所取袋子中随机取一个球，结果是白球。求"这个白球是来自1号袋子"的概率。

41 设 $A=$"诊断某人患有W病"

$B=$"某人确实患有W病"

若已知 $P(A|B)=0.95$。$P(\overline{A}|\overline{B})=0.90$，$P(B)=0.0004$。现在假定某人已被诊断患有W病，求该人确实患有W病"的概率。

42 在射击室里有9支枪，其中经试射的有2支，试射过的枪的命中率是0.8，未试射过的枪的命中率是0.1。今从射击室里任取一支枪，发射一次结果命中了。求"所取的枪已试射过"的概率。

43 已知100件产品中有10件是绝对可靠的正品，每次使用这些正品时肯定不会发生故障，而在每次使用非正品时均有0.1的可能性发生故障。现从这100件产品中随机取出一件，若使用了n次均未发生故障，问n为多大时，才能有70%的把握认为所取的产品为正品。

44 一个坛子中有b个黑球r个红球。随机取出一个球，然后放回，并且再加进c个与所取的球同样颜色的球。若已知第二次所取的球是黑球，求"第一次所取的球也是黑球"的概率。

45 将44题中取球手续进行n次，求"前面n_1次取得黑球，后面$n_2=n-n_1$次取得红球"的概率。

46 取球手续与44题同，证明：任何一次取得黑球的概率都是$b/(b+r)$

47 对于三个事件A，B，C，若

$$P(A\cap B|C)=P(A|C)\cdot P(B|C)$$

成立，则称A与B关于条件C独立。若已知A与B关于条件C与$\overline{C}$独立且

$$P(C)=0.5,\ P(A|C)=P(B|C)=0.90,$$

$$P(A|\overline{C})=0.2, P(B|\overline{C})=0.1,$$

试求

$$P(A),\ P(B),\ P(A\cap B)$$

并证明A与B不独立。

48 制造一种产品需要经过三道工序，在每道工序中出废品的概率都是0.02，如每道工序出废品与否互不影响，求“不出废品”的概率。

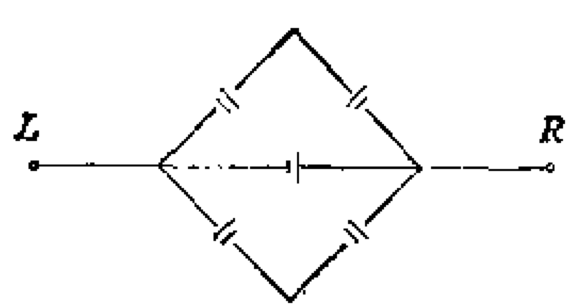
49 设A与B独立，若已知 $P(A\cup B)=0.6$，$P(A)=0.4$，求$P(B)$。

50 两个人，每个人扔三个硬币，求“二人出现正面的次数相等”的概率。

51 考虑由0和1组成的二进数（例如1011，1100等等）。假定二进数由n个数码组成，假定在每个数码上出错的概率独立地均等于p，求“一个二进数是错误的”的概率。

52 如果二阶行列式的每一个元素是0或1，求“二阶行列式的值大于0”的概率（假定行列式中各位置上的数均独立地以概率$\frac{1}{2}$选0或1）。

53 已知电路图中的各开关闭合的概率均独立地为p，求“电流由L通到R”的概率。



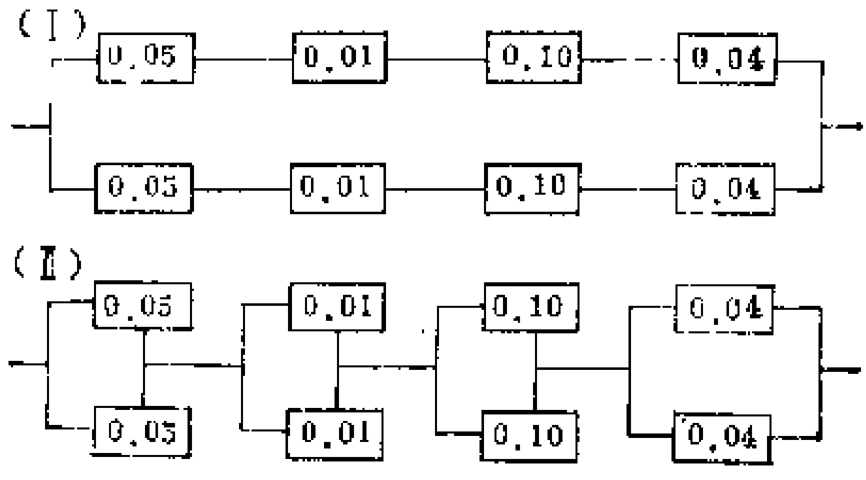
图（1—53题）

54 设敌机俯冲时被步枪击落的概率是0.008，求“当25支步枪同时开火时，击落敌机”的概率。

55 设每架高射炮命中目标的概率是0.6，今有一架敌机来侵犯，问需要多少架高射炮射击才能以0.99的概率命中它？

56 一台电子设备有两个设计方案，其框图如下：

图上标出了各元件出故障的概率。试分析两种设计哪种更可靠（即正常工作的概率大）。



图（1—56题）

学习指导

（一）本章概要

这一章主要介绍了事件及事件的概率（包括条件概率）的概念；事件的关系和运算；概率的性质；由概率的性质诱导出的概率计算公式；古典概型的概率计算方法等。这些内容是学好以后各章的基础。

这一章中的主要概念，比如，事件、事件的概率及事件的独立性等，都是先给出一个粗略的比较直观的描述，然后经过分析进一步给出严格的数学定义，这是为了使读者便于理解定义的实际涵义，不致于在一开始就由于抽象的定义而给学习带来困难。但是，请读者在学习时，一定要牢记有关概念的严格数学定义，只有一个描述性的理解是不够的，因为它不能做为严格的理论推导的基础。

这一章中古典概率的例题及习题占有很大篇幅，在学习这一章时，可多用一些工夫在解古典概型的题目上，这是因为古典概型的概率计算问题概括了很多实际问题，有很广泛的应用，同时通过解题可以初步培养解题技能，可以从各种不同角度启发读者加深对基本概念和基本内容的理解和掌握。

（二）内容说明

1. 关于事件的关系和运算

（1）事件等式的证明

由于事件是基本事件空间Ω的子集，所以，事件的运算与集合的运算完全一致，也具有下面的重要性质。

（a）交换律：$A\cup B=B\cup A$，$A\cap B=B\cap A$

（b）结合律：$(A\cup B)\cup C=A\cup(B\cup C)$

$$(A\cap B)\cap C=A\cap(B\cap C)$$

（c）分配律：$(A\cup B)\cap C=(A\cap C)\cup(B\cap C)$

$$(A\cap B)\cup C=(A\cup C)\cap(B\cup C)$$

（d）对偶原理：$\overline{A\cup B}=\overline{A}\cap\overline{B}$

$$\overline{A\cap B}=\overline{A}\cup\overline{B}$$

这些在集合论中已成立的事实，对事件也可用集合论中的方法来证明。证明两个事件A，B等价时最基本的方法就是：若基本事件$\omega\in A$，必导致$\omega\in B$，于是$A\subset B$，反之，若$\omega\in B$，必导致$\omega\in A$，于是$B\subset A$，因而得$A=B$。下面证明$\overline{A\cup B}=\overline{A}\cap\overline{B}$以作示范。

由于

$$\omega\in\overline{A\cup B}\Rightarrow\omega\overline{\in}A\cup B\Rightarrow\omega\overline{\in}A\text{且}\omega\overline{\in}B\Rightarrow\omega\in\overline{A}\text{且}\omega\in\overline{B}$$

$$\Rightarrow\omega\in\overline{A}\cap\overline{B}$$

所以 $\overline{A\cup B}\subset\overline{A}\cap\overline{B}$

由于

$$\omega\in\overline{A}\cap\overline{B}\Rightarrow\omega\in\overline{A}\text{且}\omega\in\overline{B}$$

$$\Rightarrow\omega\overline{\in}A\text{且}\omega\overline{\in}A\Rightarrow\omega\overline{\in}A\cup B$$

$$\Rightarrow\omega\in\overline{A\cup B}$$

所以 $\overline{A}\cup\overline{B}\subset\overline{A\cup B}$

因此 $\overline{A\cup B}=\overline{A}\cap\overline{B}$

对偶原理在事件的运算中经常用到。它可以推广到有限个或可列个事件的情形，即

$$\overline{A_1 \cup A_2 \cup \cdots \cup A_n \cup \cdots} = \overline{A_1} \cap \overline{A_2} \cap \cdots \cap \overline{A_n} \cap \cdots$$

$$\overline{A_1 \cap A_2 \cap \cdots \cap A_n \cap \cdots} = \overline{A_1} \cup \overline{A_2} \cup \cdots \cup \overline{A_n} \cup \cdots$$

简记为

$$\overline{\bigcup_{i=1}^{\infty} A_i} = \bigcap_{i=1}^{\infty} \overline{A_i}$$

$$\overline{\bigcap_{i=1}^{\infty} A_i} = \bigcup_{i=1}^{\infty} \overline{A_i}$$

用话来说，就是：事件并的对立事件等于对立事件的交；事件交的对立事件等于对立事件的并。

（2）关于逆事件

已知一个事件，它的逆事件是什么？在有些情况下是很容易回答的，比如，扔一枚硬币，“出正面”的逆事件是“出反面”．但在有些情况下，对初学者来说，还是要认真想一想的．比如，从0，1，2，…，9 这十个数码中随机而可重复地取出 5 个数码，事件 A = “5 个数码中至少有两个相同”的逆事件$\overline{A}$是什么呢？$\overline{A}$ = “5 个数码全不相同”．这是因为$\overline{A}$应当是由不包含在A中的基本事件所组成．这个试验的结果凡属“两个数码相同”、“3 个数码相同”、“4 个数码相同”、“5 个数码相同”的情况都包含在A中，而不包含在A中的情况都是“5 个数码全不相同”．若事件 B = “5 个数码至多有两个相同”，则它的逆事件$\overline{B}$ = “5 个数码至少有 3 个相同．”

再比如，袋中放有红球与白球，从中随机取出（有放回地）两球．A = “两球都是白球”的逆事件$\overline{A}$ = “两个球中至少有一个不是白球”，而不是“两球都是红球”．

（3）互逆与互斥

两个事件A，B，若满足两个关系（a）$A \cap B = V$，（b）$A \cup B = U$，则称其为互逆；若只满足（a），则称其为互斥．所以，互逆一定互斥，但互斥不一定互逆．比如，一批产品中有一级品、二

级品、三级品，从中随机取一件．令A＝"为一级品"，B＝"为二级品"，则A与B互斥，但不互逆，A的逆$\bar{A}$＝"不是一级品"＝"二级品或三级品"．

（4）互斥与独立

有的初学者对这两个概念弄不清，有时把它们当成一回事，这是错误的．事实上，若A与B互斥，则A与B一定不独立，这是因为A与B中的一个发生与否直接影响另一个是否发生，即A发生就决定了B一定不发生；B发生就决定了A不发生，因此，A与B互相影响而不是相互独立．另外，若A与B独立，则A与B一定不互斥，这是因为相互独立的事件决不会互相排斥．

2．关于概率的定义

在讲义中关于概率定义我们提到了古典概率定义、几何概率定义、概率的统计定义，最后提出了概率的公理化定义．这样从概率概念的发展进程来引出概率的公理化定义是为了帮助读者更好地理解概率定义的实质．现在再把这些定义重述如下．

（1）古典概率定义

对于古典型试验，若试验的一切基本事件有n个，事件A包含的基本事件有k个，则A的概率定义为

$$P(A)=\frac{A\text{包含的基本事件数}}{\text{一切基本事件数}}=\frac{k}{n}$$

（2）几何概率定义

对于几何型试验，若全空间为Ω，区域$g\subset\Omega$，则事件A＝"点M落在g中"的概率定义为

$$P(A)=\frac{g\text{的测度}}{\Omega\text{的测度}}$$

（3）概率的统计定义

当试验次数n充分大时，事件的频率总围绕某一常数摆动，这个常数就定义为该事件的概率．

（4）公理化定义

设Ω为基本事件空间，$\mathscr{F}$为关于Ω的一个事件域，如果定义在事件域$\mathscr{F}$上的一个集函数$P(A)(A\in\mathscr{F})$满足下面三个条件，则P叫做概率：

（a）$P(A)\geqslant 0$　　对一切 $A\in\mathscr{F}$

（b）$P(\Omega)=1$

（c）若 $A_i\in\mathscr{F}$，$i=1,2,\cdots$，且互不相容，则

$$P\left(\sum_{i=1}^{\infty}A_i\right)=\sum_{i=1}^{\infty}P(A_i)$$

概率的公理化定义，对古典型试验来说它与概率的古典定义一致，对几何型试验来说它与概率的几何定义一致，对一般情况来说它与概率的统计定义一致，它是统计定义的数学化。

首先，我们来说明，对古典型试验，如何由公理化定义导出概率的计算公式。

设基本事件空间由有限个基本事件 $\omega_1,\omega_2,\cdots,\omega_n$所组成，而且$P(\omega_1)=P(\omega_2)=\cdots=P(\omega_n)$，由概率定义中的（b），（c）则有

$$\begin{aligned}P(\Omega)&=P(\omega_1+\omega_2+\cdots+\omega_n)\\&=P(\omega_1)+P(\omega_2)+\cdots+P(\omega_n)=1\end{aligned}$$

所以，$P(\omega_i)=\dfrac{1}{n}$，$i=1,2,\cdots,n$。若事件A包含基本事件 ω_{i1}，$\omega_{i2},\cdots,\omega_{ik}$，即 $A=\omega_{i1}+\omega_{i2}+\cdots+\omega_{ik}$，则由概率的有限可加性（可由概率的可列可加性导出）便得

$$P(A)=P(\omega_{i1})+P(\omega_{i2})+\cdots+P(\omega_{ik})=k\cdot\frac{1}{n}=\frac{k}{n}$$

其次，对于几何型试验来说，我们所定义的几何概率，因为它满足公理化定义中概率的三条公理，所以，它与概率的公理化定义是一致的。

至于概率的统计定义与公理化定义的一致性，将由第四章中贝努利大数定理得到说明。

3．关于古典概率计算公式的应用

（1）古典概率计算公式要求试验的一切基本结果是等可能的，否则不能用，这一点要特别注意。比如，扔两枚硬币，求事件 A = “两枚都出正面”的概率。这时就不能考虑基本事件空间 Ω = {（两个正面），（一正一反），（两个反面）}，它是把两枚硬币看作是没有区别的，它的着眼点是两枚硬币中出正面的个数。这三个基本事件不是等可能的，不能认为这三个基本事件的概率都是1/3。为了应用古典概率计算公式，应当考虑基本事件空间 Ω = {(正，正)，（正，反），（反，正），（反，反)}，这四个基本事件可认为是等可能的，概率均为1/4。因此，所求概率 $P(A)$ = 1/4。

（2）应用古典概率计算公式求概率时，可以在不同的基本事件空间里考虑，只要它们都符合古典概型的条件就可以，因此，同一问题可以有不同解法。

例如，袋中有 a 个黑球，b 个白球，它们除颜色不同外，其它方面没有差别，现在随机地把球一个接一个地摸出来，求 A = “第 k 次摸出的球是黑球”的概率。

可以有两种解法。

（a）我们可以把 a 个黑球及 b 个白球都看作是不同的（比如，设想它们都编上了号），这时基本事件空间就是由 $(a+b)!$ 个基本事件所组成，它们可认为是等可能的。而 A 包含的基本事件有 $(a+b-1)!\times a$ 个，这是因为在第 k 次摸得黑球有 a 种可能情况，而另外 $(a+b-1)$ 次摸球有 $(a+b-1)!$ 种可能情况。所以，所求概率为

$$P(A)=\frac{a\times(a+b-1)!}{(a+b)!}=\frac{a}{a+b}$$

（b）依题意，黑球与黑球，白球与白球是没有区别的，所以，基本事件空间由 $\binom{a+b}{a}$ 个基本事件所组成，它们可认为是等可能的。而 A 包含的基本事件有 $\binom{a+b-1}{a-1}$ 个．这是因为第 k 次摸

得黑球相当于在$a+b$个位置上的第k个位置是黑球，而其余$a+b-1$个位置上可任取$a-1$个位置放黑球．所以，所求概率为

$$P(A)=\frac{\binom{a+b-1}{a-1}}{\binom{a+b}{a}}=\frac{a}{a+b}$$

从这个例子可以看出，对于同一随机现象可以用不同的模型来描述．但是，在用古典概率计算公式时，分母、分子的计算必须在同一基本事件空间中来考虑．如果考虑顺序，则分母、分子都要考虑顺序，如果不考虑顺序，则分母、分子都不考虑顺序．

（3）超几何分布

在用古典概率公式求概率的问题中有一类是属于超几何分布的，这在讲义§1.3中已举了好几个例子．这里要说明的是，可以把这类问题进一步推广，推广到由N个元素组成的总体中含有好几种不同元素的情况中去．比如，假设总体中含有三种元素，其个数分别为$M_1,M_2,N-M_1-M_2$．如果从总体中不放回地取出n个元素，则$A=$“n个元素中含有m_1个第一种元素，m_2个第二种元素，$n-m_1-m_2$个第三种元素”的概率应为

$$P(A)=\frac{\binom{M_1}{m_1}\binom{M_2}{m_2}\binom{N-M_1-M_2}{n-m_1-m_2}}{\binom{N}{n}}$$

这里要求 $m_1\leqslant M_1$，$m_2\leqslant M_2$，$n-m_1-m_2\leqslant N-M_1-M_2$．它的导出与例1,3,2类似．

例 1.1　每一页书中都有N个印刷符号可能误印，全书共有n页，求$A=$“全书共有r个印错的符号中，第1、第2、…，第n页分别含有$r_1,r_2,\cdots r_n$个符号误印”的概率，$r=r_1+r_2+\cdots+r_n$．答案是

$$P(A)=\frac{\binom{N}{r_1}\binom{N}{r_2}\cdots\binom{N}{r_n}}{\binom{nN}{r}}$$

4．关于几何概率

比较古典概率定义与几何概率定义可以看到，它们的区别只不过是把古典概率公式中Ω中的基本事件的总数n，换成了Ω的测度，而把A所包含的基本事件数k，换成了A所包含的基本事件构成的区域g的测度．试验的一切结果由有限等可能推广到无限等可能，必然这样引出几何概率定义。

下面再举两个几何概率的例子

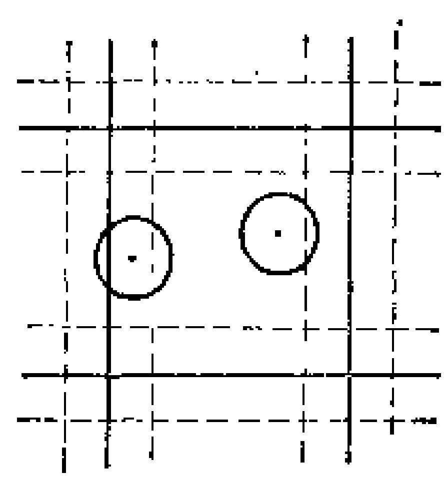
例 1.2　在方格纸上横竖引间隔为8cm的直线，于是做成很多边长为8cm的正方形．将直径为3cm的圆板随机投上，求下面事件的概率．

$A=$ “圆板落入正方形中，

$B=$ “圆板与正方形的边相交，

$C=$ “圆板跨四个正方形．

解　如图（1—1）所示，圆板中心位置的一切可能对应边长为8cm的正方形．而事件A发生的充要条件是圆板中心位置在边长为5cm的正方形中；$B=\overline{A}$；事件C发生的充要条件是圆板中心位置在边长为3cm的正方形中．所以，所求概率为



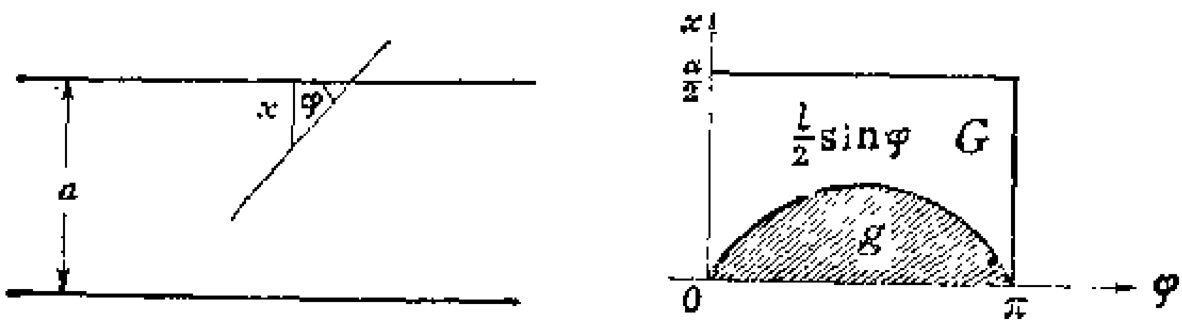
图（1—1）

$$P(A)=\frac{25}{64}\quad P(B)=\frac{39}{64}$$

$$P(C)=\frac{9}{64}$$

例 1.3　蒲丰(*Buffon*)问题　平面上画有一些等距离的平行线，它们之间的距离都为a，向此平面随机投一长度为l $(l<a)$的针，求$A=$ “针与一平行线相交”的概率．

解　用x表示针的中点到最近一条平行线的距离，用φ表示针与此线的交角（图1—2），显然有



图（1—2）

$$0\leqslant x\leqslant\frac{a}{2},\quad 0\leqslant\varphi\leqslant\pi$$

这两个式子决定了 $\varphi-x$ 平面上一矩形 G，它对应投针的一切可能结果．而针与一平行线相交的充要条件是 $x\leqslant\frac{l}{2}\sin\varphi$，这个不等式决定了 G 中的一个区域 g，它对应事件 A 包含的结果．所以，所求概率为

$$P(A)=\frac{g\text{的面积}}{G\text{的面积}}=\frac{\frac{1}{2}\int_0^\pi l\sin\varphi d\varphi}{\frac{1}{2}a\pi}=\frac{2l}{\pi a}$$

由于 $P(A)$ 与 π 有关，因此可利用它们的关系求 π 的近似值．方法是这样：投针多次，比如 N 次，计算针与平行线相交的次数 n，用频率 $\frac{n}{N}$ 近似代替概率 $P(A)$，则由 $P(A)=\frac{2l}{\pi a}$，可得 π 的近似值 $\frac{2lN}{an}$．

5．关于概率性质的证明

（1）在证明概率性质时，有几个性质都是这样证明的，即，把一个事件分解为两个互斥事件的和，于是该事件的概率就变成这两个互斥事件概率的和，从面求得要证明的结果．这是证明概率等式时常用的方法．

比如，证明性质 4（若 $A\subset B$，则 $P(B)\geqslant P(A)$）时，是先把事件 B 分解为两个互斥事件 A 与 $(B-A)$ 的和，于是 $P(B)=P(A)+P(B-A)$，所以得

$P(B)\geqslant P(A)$

再比如，证明性质（6）（$P(A\cup B)=P(A)+P(B)-P(AB)$）时，是先把 $A\cup B$ 分解为两个互斥事件 A 与 $(B-A)$ 的和，于是 $P(A\cup B)=P(A)+P(B-A)$，再把事件 B 分解为互斥事件 $B-A$ 与 AB 的和，从而得 $P(B)=P(B-A)+P(AB)$，将 $P(B-A)=P(B)-P(AB)$ 代入 $P(A\cup B)=P(A)+P(B-A)$ 中，便得性质 6。

还有性质 8 也是这样证明的，不再重复了

（2）概率一般加法公式的证明

一般加法公式　对任意 n 个事件 $A_1,A_2,\cdots,A_n$，有

$$P\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}P(A_i)-\sum_{1\leqslant i<j\leqslant n}P(A_iA_j)+\sum_{1\leqslant i<j<k\leqslant n}P(A_iA_jA_k)-\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n)$$

证明　用归纳法。已知当 $n=2$ 时公式成立（见1.5.3)），即

$$P(A_1\cup A_2)=P(A_1)+P(A_2)-P(A_1A_2)$$

今设　$n=k$ 时公式成立，则根据（1.5.3）有

$$\begin{aligned}P\left(\bigcup_{i=1}^{k+1}A_i\right)&=P\left[\left(\bigcup_{i=1}^{k}A_i\right)\cup A_{k+1}\right]\\&=P\left(\bigcup_{i=1}^{k}A_i\right)+P(A_{k+1})-P\left(\bigcup_{i=1}^{k}A_iA_{k+1}\right)\\&=\sum_{i=1}^{k}P(A_i)-\sum_{1\leqslant i<j\leqslant k}P(A_iA_j)+\cdots\\&\quad+(-1)^{k-1}P(A_1A_2\cdots A_k)+P(A_{k+1})\\&\quad-P\left(\bigcup_{i=1}^{k}A_iA_{k+1}\right)\qquad(1.1)\end{aligned}$$

令　　$B_i=A_iA_{k+1}$，则

$$P\left(\bigcup_{i=1}^{k}A_iA_{k+1}\right)=P\left(\bigcup_{i=1}^{k}B_i\right)$$

$$= \sum_{i=1}^{k} P(B_i) - \sum_{1 \le i < j \le k} P(B_i B_j) + \cdots$$

$$+ (-1)^{k-1} P(B_1 B_2 \cdots B_k)$$

$$= \sum_{i=1}^{k} P(A_i A_{k+1}) - \sum_{1 \le i < j \le k} P(A_i A_j A_{k+1}) + \cdots$$

$$+ (-1)^{k-1} P(A_1 A_2 \cdots A_k A_{k+1})$$

将上式代入（1.1）便得

$$P\left(\bigcup_{i=1}^{k+1} A_i\right) = \sum_{i=1}^{k} P(A_i) - \sum_{1 \le i < j \le k} P(A_i A_j) + \cdots$$

$$+ (-1)^{k-1} P(A_1 A_2 \cdots A_k)$$

$$+ P(A_{k+1}) - \sum_{i=1}^{k} P(A_i A_{k+1}) + \sum_{1 \le i < j \le k} P(A_i A_j A_{k+1})$$

$$- \cdots + (-1)^{k} P(A_1 A_2 \cdots A_{k+1})$$

将有关项合并即知公式当 $n=k+1$ 时成立。

下面举几个应用一般加法公式的例子。

例 1.4　将数码 $1,2,\cdots,n$ 随机排成一列，求下面两个事件的概率。

$A=$“至少有一个数码与它所占位置的号数一致”

$B=$“没有一个数码与它所占位置的号数一致”

解　先求 $P(A)$。

设　$A_i=$“第 i 个数码占有第 i 个位置”

$i=1,2,\cdots,n$，则

$$A = A_1 \cup A_2 \cup \cdots \cup A_n \qquad P(A) = P\left(\bigcup_{i=1}^{n} A_i\right)$$

n 个数码的全排列共有$n!$种排列法，而当第 i 个数码占有第 i 个位置时，其余 $n-1$ 个数码的全排列共有 $(n-1)!$ 种排列法；当第 i，j 个数码分别占有第 i，j 个位置时，其余$n-2$个数码的全排列共有$(n-2)!$ 种排列法；…。所以有

$$P(A_i) = \frac{(n-1)!}{n!} = \frac{1}{n}$$

$$P(A_iA_j) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}$$

$$\cdots\cdots$$

$$P(A_1A_2\cdots A_n) = \frac{1}{n!}$$

因而

$$P(A) = \binom{n}{1}\frac{1}{n} - \binom{n}{2}\frac{1}{n(n-1)} + \cdots + (-1)^{n-1}\frac{1}{n!}$$

$$= 1 - \frac{1}{2!} + \frac{1}{3!} - \cdots + (-1)^{n-1}\frac{1}{n!}$$

$$= \sum_{k=1}^{n}\frac{(-1)^{k-1}}{k!}$$

其次求 $P(B)$

$$P(B) = 1 - \sum_{k=1}^{n}\frac{(-1)^{k-1}}{k!} = \sum_{k=2}^{n}\frac{(-1)^k}{k!}$$

$$= \sum_{k=0}^{n}\frac{(-1)^k}{k!}$$

有许多表面上提法不同的问题，实际上属于这同一类型，比如下面两例。

例 1.5　某班有 n 个战士，每人各有一支枪，这些枪外形完全一样，在一次夜间紧急集合中，每人都随机地取走一支枪，求下面两个事件的概率，A = "至少有一个战士拿到自己的枪，B = "n 个战士都没有拿到自己的枪"．

例 1.6　有 n 个写好地址的信封，还有与其对应的 n 封信．现在将 n 封信随机放入 n 个信封中，求下面两个事件的概率，A = "至少有一封信的内容与地址一致"，B = "没有一封信的内容与地址一致"．

6．关于事件和的概率的计算

综合所学的概率性质，关于若干个事件和的概率的计算可归纳为以下三种情况．

（1）若 $A_1, A_2, \cdots, A_n$ 互斥，则 $A=\sum_{i=1}^{n} A_i$ 的概率为

$$P(A)=\sum_{i=1}^{n} P(A_i)$$

（2）若 $A_1, A_2, \cdots, A_n$ 相互独立，则可把求事件和的概率转化为求独立事件积的概率。

由于 $A=\bigcup_{i=1}^{n} A_i$ 的逆事件为 $\overline{A}=\bigcap_{i=1}^{n}\overline{A_i}$，所以

$$P(A)=1-P(\overline{A})=1-P\left(\bigcap_{i=1}^{n}\overline{A_i}\right)=1-\prod_{i=1}^{n} P(\overline{A_i})$$

（3）若 $A_1, A_2, \cdots, A_n$ 既不互斥也不相互独立，则可用一般加法公式（对任意 n 个事件和，都可用一般加法公式）。

$$P(A)=\sum_{i=1}^{n} P(A_i)-\sum_{1\leqslant i<j\leqslant n} P(A_iA_j)$$
$$+\sum_{1\leqslant i<j<k\leqslant n}(A_iA_jA_k)$$
$$-\cdots+(-1)^{n-1}P(A_1A_2\cdots A_n)$$

7. 关于全概率公式与贝叶斯公式

这两个公式是概率计算的两个重要公式，从形式上看好像很复杂，其实导出它们的想法是很容易理解的。如果能从它们的导出来理解它们，不仅容易记住它们而且还会应用。

（1）全概率公式

假如有有限个或可列个事件 $A_1, A_2, \cdots, A_n, \cdots$，它们互不相容，且 $\sum_i A_i=U$，即每次试验必发生其一。这时，我们所考虑的事件 B，它的发生就一定与 $A_1, A_2, \cdots, A_n$ 中的一个同时发生，这样，我们就可以把事件 B 分解成若干个简单的互斥事件的和，而求 B 的概率就变成求这些简单事件的概率的和。用式子表示就是

$$P(B)=P(BU)=P\left[B\left(\sum_i A_i\right)\right]=P\left(\sum_i BA_i\right)$$

$$=\sum_i P(BA_i)$$

接下去就是如何求各 $P(BA_i)$，这只要应用乘法公式就行了．于是得

$$P(B)=\sum_i P(A_i)P(B|A_i)$$

在实际问题中，当我们要求某一事件 B 的概率时，是否存在满足全概率公式的 $A_1,A_2,\cdots$，它们具体内容是什么，这要具体问题具体分析．下面举一个例子．

例 1.7 在一个袋中放有 m 张写有数码 1 的小纸片和 n 张写有数码 2 的小纸片，另外，在坛子 Ⅰ 中放有 r_1 个红球和 b_1 个黑球；在坛子 Ⅱ 中放有 r_2 个红球和 b_2 个黑球．现在从袋中随机取一纸片，若其数码为 1，则从坛子 Ⅰ 中取一球，若其数码为 2，则从坛子 Ⅱ 中取一球，求所取球为红球的概率．

解 令 $B=$"所取球为红球"．显然，取球结果与从袋中取出的数码有关，而从袋中取出的数码有两种可能．令

$A_1=$"取得数码 1"

$A_2=$"取得数码 2"

由于 $A_1\cup A_2=U$，$A_1\cap A_2=V$，所以

$$B=BA_1+BA_2$$

$$P(B)=P(BA_1)+P(BA_2)$$

$$=P(A_1)P(B|A_1)+P(A_2)P(B|A_2)$$

而由题意可知 $P(A_1)=\dfrac{m}{m+n}$，$P(A_2)=\dfrac{n}{m+n}$；

$P(B|A_1)=\dfrac{r_1}{r_1+b_1}$，$P(B|A_2)=\dfrac{r_2}{r_2+b_2}$

所以

$$P(B)=\frac{m}{m+n}\cdot\frac{r_1}{r_1+b_1}+\frac{n}{m+n}\cdot\frac{r_2}{r_2+b_2}$$

这里，我们没有直接用公式，而是按导出公式的顺序来解此题的，这是为了使读者熟悉导出公式的过程，并不是说解此类题时都要这样做，对已熟悉公式的读者，可直接将有关数值代入公式进行计算。

（2）贝叶斯公式

假如有有限个或可列个事件 $A_1, A_2, \cdots$，它们互不相容，且 $\sum_i A_i = U$，它们是导致某一事件 B 发生的可能原因。如果在一次试验中事件 B 发生了，这时要求"B 的发生是由 A_i 所引起"的概率是多少，即求 $P(A_i|B)$，要用贝叶斯公式。

只要记住条件概率定义，贝叶斯公式就不难导出，按条件概率定义有

$$P(A_i|B) = \frac{P(A_iB)}{P(B)}$$

然后分子、分母分别用乘法公式及全概率公式，就得到了贝叶斯公式

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_i P(A_i)P(B|A_i)}$$

贝叶斯公式很有用，有时根据已有的经验或历史统计资料，我们可以对 $P(A_i)$ 做出一个估计，而且也可以知道 $P(B|A_i)$。现在如果我们又有了一个新的情况，即事件 B 发生了，这时，这个新的信息，补充了我们对 A_i 发生可能性的认识，这样求得的 A_i 的条件概率 $P(A_i|B)$ 既利用了历史经验资料，又把新得到的情况考虑进去了，因此，它给出的结果可能更符合实际。这在数理统计中是很有用的。

例 1.8　已知具有某种症状的患者为癌患者的概率为0.05，在有该症状的患者中，癌患者经仪器检查为阳性的概率为0.85；非癌患者经仪器检查为阳性的概率为0.05。现有一患者谈了他具有该症状、且经仪器检查为阳性，求该患者为癌患者的概率。

解　由题意可知，所讨论的问题完全限制在具有某种症状的患者中．在这个范围内，我们所求的概率为条件概率 P（为癌患者｜仪器检查为阳性）．令

$B=$ “仪器检查为阳性”

$A=$ “为癌患者”

$\overline{A}=$ “为非癌患者．

则所求概率为

$$P(A|B)=\frac{P(AB)}{P(B)}=\frac{P(A)P(B|A)}{P(A)P(B|A)+P(\overline{A})P(B|\overline{A})}$$

$$=\frac{0.05\times0.85}{0.05\times0.85+0.95\times0.05}$$

$$=0.472$$

习 题 解 答

1　（1）$\Omega=\{\omega_1,\omega_2,\omega_3,\omega_4,\omega_5,\omega_6,\omega_7,\omega_8\}$

$\omega_1=$（正正正）　　$\omega_2=$（正正反）

$\omega_3=$（正反正）　　$\omega_4=$（反正正）

$\omega_5=$（反反正）　　$\omega_6=$（反正反）

$\omega_7=$（正反反）　　$\omega_8=$（反反反）

（2）$\Omega=\{\omega_1,\omega_2,\omega_3,\cdots,\omega_{10}\}$

$\omega_1=\{白_1\}$　$\omega_2=\{白_2\}$　$\omega_3=\{白_3\}$　$\omega_4=\{白_4\}$

$\omega_5=\{红_1\}$　$\omega_6=\{红_2\}$　$\omega_7=\{红_3\}$　$\omega_8=\{红_4\}$

$\omega_9=\{红_5\}$　$\omega_{10}=\{红_6\}$

2　（1）$A\cap\overline{B}=\{\omega_1,\omega_2,\omega_3\}$

（2）$(A\cap\overline{B})\cup(\overline{A}\cap B)=\{\omega_1,\omega_2,\omega_3,\omega_7,\omega_8,\omega_9\}$

（3）$\overline{A\cup B}=\{\omega_{10},\omega_{11},\omega_{12}\}$

（4）$\overline{A}\cap\overline{B}=\{\omega_{10},\omega_{11},\omega_{12}\}$

（5）$A\cap B=\{\omega_1,\omega_2,\omega_3,\omega_7,\omega_8,\omega_9,\omega_{10},\omega_{11},\omega_{12}\}$

（6） $\overline{A}\cup\overline{B}=\{\omega_1,\omega_2,\omega_3,\omega_7,\omega_8,\omega_9,\omega_{10},\omega_{11},\omega_{12}\}$

3　$\overline{A}$ = “三件全是好品” $\overline{A}=C$

$\overline{B}$ = “三件中至少有两件好品”

= “三件中至多有一件废品”

$A\cup C=U$（必然事件）

$A\cap C=V$（不可能事件）

$A-B=$ “三件中恰有一件废品”

4　（1） $A\overline{B}\,\overline{C}$

（2） $\overline{A}BC$，

（3） ABC，

（4） $\overline{A}\,\overline{B}\,\overline{C}$，

（5） $A\overline{B}\,\overline{C}\cup\overline{A}B\overline{C}\cup\overline{A}\,\overline{B}C$，

（6） $AB\overline{C}\cup A\overline{B}C\cup\overline{A}BC$，

（7） $A\cup B\cup C$，

（8） $AB\cup BC\cup AC$，

（9） $\overline{ABC}$。

5　证明两个事件 A，B 等价（$A=B$）的最基本方法是用等价的定义，即证明如下的两个包含关系成立。

$A\subset B$，$B\subset A$

（1）见第一章学习指导。

（2） $\overline{A\cap B}=\overline{A}\cup\overline{B}$

由于　$\omega\in\overline{A\cap B}\Rightarrow\omega\overline{\in}A\cap B\Rightarrow\omega\overline{\in}A$ 或 $\omega\overline{\in}B\Rightarrow\omega\in\overline{A}$ 或 $\omega\in\overline{B}$

$\Rightarrow\omega\in\overline{A}\cup\overline{B}$，所以，

$\overline{A\cap B}\subset\overline{A}\cup\overline{B}$　　（i）

由于　$\omega\in\overline{A}\cup\overline{B}\Rightarrow\omega\in\overline{A}$ 或 $\omega\in\overline{B}\Rightarrow\omega\overline{\in}A$ 或 $\omega\overline{\in}B$

$\Rightarrow\omega\overline{\in}A\cap B\Rightarrow\omega\in\overline{A\cap B}$，所以，

$\overline{A}\cup\overline{B}\subset\overline{A\cap B}$　　（ii）

结合（i）和（ii），则得：$\overline{A\cap B}=\overline{A}\cup\overline{B}$。

（3） $(A\cup B)-B=A\cap\overline{B}$

由于 $\omega\in(A\cup B)-B\Rightarrow\omega\in(A\cup B)$且 $\omega\overline{\in}B\Rightarrow\omega\in A$且$\omega\in\overline{B}$

$\Rightarrow\omega\in A\cap\overline{B}$所以

$$(A\cup B)-B\subset A\cap\overline{B} \qquad \text{(i)}$$

又由于 $\omega\in A\cap\overline{B}\Rightarrow\omega\in A$且$\omega\overline{\in}B\Rightarrow\omega\in(A\cup B)$且$\omega\overline{\in}B$

$\Rightarrow\omega\in(A\cup B)-B$ 所以

$$A\cap\overline{B}\subset(A\cup B)-B \qquad \text{(ii)}$$

结合(i)和(ii)，便得 $(A\cup B)-B=A\cap\overline{B}$

（4）$(A\cup B)-AB=A\overline{B}\cup\overline{A}B$

由于 $\omega\in[(A\cup B)-AB]\Rightarrow\omega\in(A\cup B)$且$\omega\overline{\in}AB$

$\Rightarrow\omega\in A$、$\omega\overline{\in}AB$ 或 $\omega\in B$、$\omega\overline{\in}AB$

$\Rightarrow\omega\in A$、$\omega\overline{\in}B$ 或 $\omega\in B$、$\omega\overline{\in}A$

$\Rightarrow\omega\in A\overline{B}$ 或 $\omega\in B\overline{A}\Rightarrow\omega\in A\overline{B}\cup B\overline{A}$，所以

$$(A\cup B)-AB\subset A\overline{B}\cup\overline{A}B \qquad \text{(i)}$$

又由于 $\omega\in A\overline{B}\cup\overline{A}B\Rightarrow\omega\in A\overline{B}$ 或 $\omega\in\overline{A}B$

$\Rightarrow\omega\in A$，$\omega\overline{\in}B$ 或 $\omega\in B$，$\omega\overline{\in}A$

$\Rightarrow\omega\in A$，$\omega\overline{\in}AB$ 或 $\omega\in B$，$\omega\overline{\in}AB$

$\Rightarrow\omega\in(A\cup B)$ 且 $\omega\overline{\in}AB$

$\Rightarrow\omega\in[(A\cup B)-AB]$ 所以

$$A\overline{B}\cup\overline{A}B\subset(A\cup B)-AB \qquad \text{(ii)}$$

结合(i)和(ii)，便得$(A\cup B)-AB=A\overline{B}\cup\overline{A}B$

（5）略

6 $A\triangleq$“恰有3粒虫豆”，此题与例1.3.2完全类似，是一个超几何分布的问题，所求概率为

$$P(A)=\frac{\binom{10}{3}\binom{90}{2}}{\binom{100}{5}}=0.0063835$$

• $A\triangleq B$ 表示令 $A=B$

7　$A \triangleq$ “三件中至少有一件不合格品”

先考虑 $\overline{A}$ = “三件全是合格品”，

$$P(\overline{A}) = \frac{\binom{46}{3}}{\binom{50}{3}} \doteq 0.77449$$

所以，　$P(A) = 1 - P(\overline{A}) \doteq 1 - 0.77449 = 0.22551$

8　$A \triangleq$ “取得两个红球和三个白球”

由于取球是有放回的，所以基本事件空间含有 12^5 个基本事件。事件 A 包含 $\binom{5}{2}4^2 \cdot 3^3$ 个基本事件，这是因为取得两个红球有 4^2 种方法，取得三个白球有 3^3 种方法，而两次取得红球在五次抽球中有 $\binom{5}{2}$ 种不同的情况。所以，

$$P(A) = \frac{\binom{5}{2} \cdot 4^2 \cdot 3^3}{12^5} = 0.017361$$

9　$A \triangleq$ “两组中男生、女生各一半”

此题与例1.3.3类似，用与例1.3.3相同的方法得

$$P = \frac{\binom{18}{9}\binom{18}{9}}{\binom{36}{18}} = 0.26048$$

10　$A_1 \triangleq$ “三个中最小数码是5”

$A_2 \triangleq$ “三个中最大数码是5”

从10个数码中任取3个，基本事件的总数为 $\binom{10}{3}$。A_1 包含 $\binom{5}{2}$ 个基本事件，这是因为三个数码中最小的数码是5，其余2个应该从6,7,8,9,10这五个数码中任取。于是，

$$P(A_1)=\frac{\binom{5}{2}}{\binom{10}{3}}=\frac{5\times4\times3}{10\times9\times8}=\frac{1}{12}$$

A_2 包含$\binom{4}{2}$个基本事件，这是因为三个数码中最大的数码是5，其余两个应该从1，2，3，4这四个数码中选取．所以，

$$P(A_2)=\frac{\binom{4}{2}}{\binom{10}{3}}=\frac{4\times3\times3}{10\times9\times8}=\frac{1}{20}$$

11　$A\triangleq$“k个数码中既无0又无1”

由于取数码是可重复的，所以基本事件空间由10^k个基本事件组成．A中包含8^k个基本事件，这是因为要求k个数码中既无0又无1，则这k个数码是从2，3，4，5，6，7，8，9这八个数码中有重复地选取的．所以，

$$P(A)=\frac{8^k}{10^k}=0.8^k$$

12　$A\triangleq$“至少有两名学生的生日在同一月”，它的对立事件$\overline{A}$为“n个学生的生日的所在月份全不相同”．先求$P(\overline{A})$．

由于每一个学生的生日的所在月份均可能是1月，2月，…，12月这十二个月中的任一个．因此，n个学生的生日所在月的基本的情况数为12^n（基本事件空间所包含的样本点的个数），而要求n个学生的生日所在月全不相同，则这n个月应该从1，2，…，12中选取，所以A包含${}_{12}P_n$个基本事件．这样一来，

$$P(\overline{A})=\frac{{}_{12}P_n}{12^n}$$

所以

$$P(A)=1-P(\overline{A})=1-\frac{{}_{12}P_n}{12^n}$$

13　$A_1\triangleq$“第一次取得奇数”

$A_2 \triangleq$ "第二次取得奇数"

$A_3 \triangleq$ "两次都取得奇数"

此题即往两张有顺序的空白卡片上不重复地随机写数码（1，2，3，4，5中之一）的问题。

第一张卡片有5种写法，相应于它的1种写法，第二张卡片有4种写法，因此，所有的基本事件的总数为$5\cdot4=20$。

A_1 包含的基本事件数为$3\cdot4=12$。这是因为第一张卡片写奇数，可写1，3，5这三个数码之一，即有三种写法，相应于它的一种写法，第二张卡片还可写上其余的四个数码中的一个。所以，

$$P(A_1)=\frac{3\cdot4}{5\cdot4}=\frac{3}{5}$$

同样的理由，第二张卡片写奇数，有三种写法，相应于它的一种写法，第一张卡片有四种写法，因此，A_2 包含$4\cdot3$个基本事件。所以，

$$P(A_2)=\frac{4\cdot3}{5\cdot4}=\frac{3}{5}$$

A_3 包含$3\cdot2$个基本事件，因此，

$$P(A_3)=\frac{3\cdot2}{5\cdot4}=\frac{3}{10}$$

14　$A \triangleq$ "至少有一个数码与它所占的位置的号数一致"，先求 $\overline{A}$ = "三个数码与它所占的位置的号数全不一致"的概率。

将1,2,3三个数码随机排成一列，有$3!$种不同的排法，因此共有$3!=6$个基本事件。$\overline{A}$包含两个基本事件（3,1,2），（2,3,1），所以，

$$P(\overline{A})=\frac{2}{6}=\frac{1}{3}$$

于是　$P(A)=1-P(\overline{A})=\frac{2}{3}$

15　$A\triangleq$“两纸片上的数的和等于10”

从 1,2,…,10 这10个数中任取两个，有$\binom{10}{2}$种不同的方法，因此基本事件的总数为$\binom{10}{2}$.

A中包含4个基本事件，即：(1,9)，(2,8)，(3,7)，(4,6)。所以，

$$P(A)=\frac{4}{\binom{10}{2}}=\frac{4}{45}$$

16　$A\triangleq$“能被 6 或 8 整除”

基本事件空间包含50个基本事件，A中包含12个基本事件，即能被 6 或 8 整除的数有12个：6，8，12，16，18，24，30，32,36,40,42,48。所以

$$P(A)=12/50=0.24$$

17　$A\triangleq$“两个数是相连的数”

在不放回的情况，所有的基本事件有$n(n-1)$个。A中含有$2(n-1)$个基本事件，即

(1,2)，(2,3),…$(n-1,n)$，　(2,1)，(3,2),…，$(n,n-1)$。

所以，

$$P(A)=\frac{2(n-1)}{n(n-1)}=\frac{2}{n}$$

在有放回的情况，所有的基本事件有n^2个，A中所含的基本事件数仍为 $2(n-1)$。因此，

$$P(A)=2(n-1)/n^2$$

18　$A\triangleq$“恰在第 k 次试验时打开门”

实际上只进行 k 次试验，一旦在第 k 次打开门，任何一个人也不会再继续去试开，由此基本事件空间包含$n(n-1)\cdots(n-k+1)$个基本事件。恰在第 k 次打开门，意味着第一次试验取出的是不能开

门的$n-1$把钥匙之一，第二次试验取出的是其余$n-2$把不能开门的钥匙之一，如此等等，因而A中含有的基本事件数为

$$(n-1)(n-2)\cdots(n-k+1)\cdot 1$$

所以，

$$P(A)=\frac{(n-1)(n-2)\cdots(n-k+1)}{n(n-1)\cdots(n-k+1)}=\frac{1}{n}$$

19　$A_1 \triangleq$"在数列中1与2相继出现"

$A_2 \triangleq$"在数列中1,2,3相继出现"

将数$1,2,\cdots,n$按任意顺序排列，共有$n!$种不同的排法，因此，所有的基本事件数为$n!$个。1,2相继出现在1,2位置时，其余$n-2$个数仍可任意排列，有$(n-2)!$种排法，但，(1,2)可能出现在(1,2)，(2,3)，$\cdots(n-1,n)$这$n-1$个位置之一，因而，A含有$(n-1)(n-2)!$个基本事件。所以，

$$P(A_1)=\frac{(n-1)(n-2)!}{n!}=\frac{1}{n}$$

同样的计算方法，可知：

$$P(A_2)=\frac{(n-2)(n-3)!}{n!}=\frac{1}{n(n-1)}$$

这个问题也可以这样想，把(1,2)想像为一个不分开的整体，于是A_1所含的基本事件数可看做是$n-1$个不同元素（其中一个元素为(1,2)）的全排列数，应为$(n-1)!$，所以，

$$P(A_1)=\frac{(n-1)!}{n!}=\frac{1}{n}$$

也可用同样想法求$P(A_2)$。

20　$A \triangleq$"5个数码各取5次"

由于取数码是可重复的，所以，基本事件的总数为5^{25}。将25个数码分成5个有序组，有$\frac{25!}{5!5!5!5!5!}$种不同的分法，因而A中含有的基本事件数为$\frac{25!}{(5!)^5}$，所以，

$$P(A)=\frac{\frac{25!}{(5!)^5}}{5^{25}}=\frac{25!}{5^{25}\cdot(5!)^5}=0.0020917$$

$\frac{25!}{5!5!5!5!5!}$ 可以这样考虑得到，先从25个数码中取出 5 个为一组，有$\binom{25}{5}$种取法，然后再在余下的20个数码中取出 5 个为一组，有$\binom{20}{5}$种取法，再在余下的15个数码中取出 5 个为一组，有$\binom{15}{5}$种取法，最后再从余下的10个数码中取出 5 个为一组，有$\binom{10}{5}$种取法。所以，将25个数码分成 5 个有序组有

$$\binom{25}{5}\binom{20}{5}\binom{15}{5}\binom{10}{5}$$
$$=\frac{25!}{5!20!}\times\frac{20!}{5!15!}\times\frac{15!}{5!10!}\times\frac{10!}{5!5!}=\frac{25!}{5!5!5!5!5!}$$

种不同分法。

21　$A\triangleq$ “某指定的盒子是空的”

将 r 个球放入 n 个盒中有n^r种不同的放法，因此，基本事件总数为 n^r。而 A所包含的基本事件数等于将 r 个球放入 $n-1$ 个盒中的不同放法数，为$(n-1)^r$。所以

$$P(A)=\frac{(n-1)^r}{n^r}=\left(1-\frac{1}{n}\right)^r$$

22　此题类似于第12题，只不过那里考虑的是月份，而这里考虑的是生日。假定一年有365天，则所求概率为

$$P(A)=\frac{{}_{365}P_n}{365^n}$$

23.　$A\triangleq$ “第 k 次取得黑球”

由于是不放回地取球$a+b$次，所以基本事件总数为$(a+b)!$。第 k 次取得黑球可以是 a 个黑球中的任一个，有 a 种不同的取法，而其余$a+b-1$个球的不同取法有$(a+b-1)!$种，所以 A 中含有 $a(a+$

$b-1)!$个基本事件．因此，所求概率为

$$P(A)=\frac{a(a+b-1)!}{(a+b)!}=\frac{a}{a+b}$$

24　$A\triangleq$"必须的抽取次数为 r"

如果$r<k$，则抽 r 次抽得 k 个白球是不可能的，这时$P(A)=0$。

当 $r\geqslant k$ 时，必须的抽取次数为 r，意即取出的 r 个球中有 k 个白球，且第 r 个球一定是白球，而其余 $k-1$ 个白球分布在前 $r-1$ 个位置的任意 $k-1$ 个位置上，因此有 $\binom{r-1}{k-1}b^k a^{r-k}$ 种不同的取法。而基本事件总数为$(a+b)^r$，所以

$$P(A)=\frac{\binom{r-1}{k-1}b^k a^{r-k}}{(a+b)^r}$$

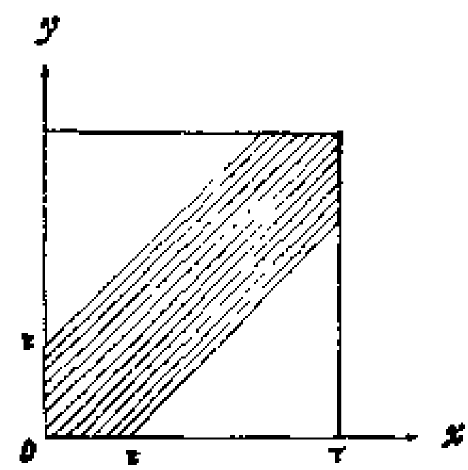
25　$A\triangleq$"点C与点O的距离$\geqslant l$"，由几何概率定义，立即可得

$$P(A)=\frac{L-l}{L}=1-\frac{l}{L}$$

26　$A\triangleq$"收音机受到干扰"

设两信号到达的时刻分别为 x，y，

由于 $0\leqslant x\leqslant T$，$0\leqslant x\leqslant T$，所以，两信号到达的一切可能结果对应边长为 T 的正方形中的一切点。而收音机受到干扰的充要条件是



图（1—26题）

$$|x-y|\leqslant\tau$$

它对应图（1—26题）中阴影部分中的全部点，所以

$$P(A)=\frac{T^2-2\cdot\frac{1}{2}(T-\tau)^2}{T^2}=1-\left(1-\frac{\tau}{T}\right)^2$$

27　$A_0\triangleq$"三线段OA、OB，OC构成三角形"，令O点为原点，

A 点坐标为 x， B 点坐标为 y， C 点坐标为 z．由于$0\leqslant x\leqslant l$，$0\leqslant y\leqslant l$，$0\leqslant z\leqslant l$，所以，A，B，C 三点的一切可能位置对应于正方体$ODEFGHIJ$中的一切点，它的体积为 l^3，而OA，OB，OC构成三角形的充要条件是

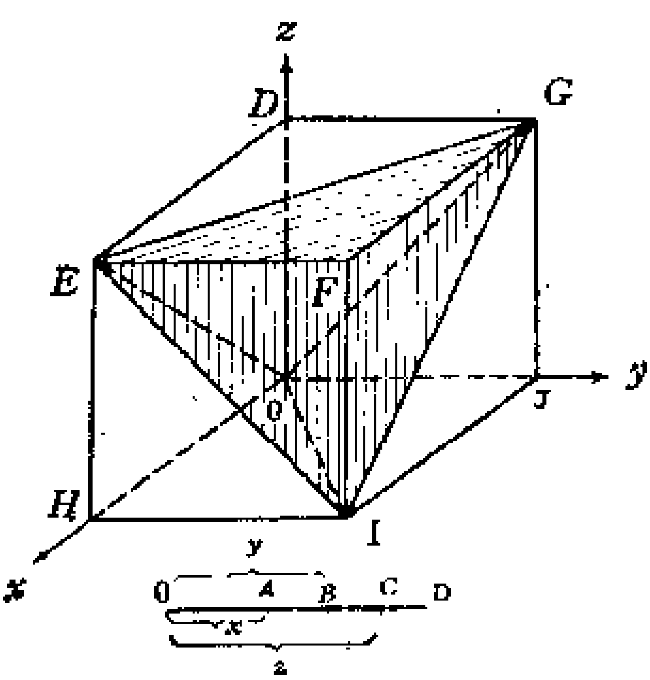
$$x+y>z,\quad x+z>y,\quad y+z>x$$

它对应图（1—27题）中正方体去掉三个三棱锥后的部分（即$OEFGI$），它的体积为

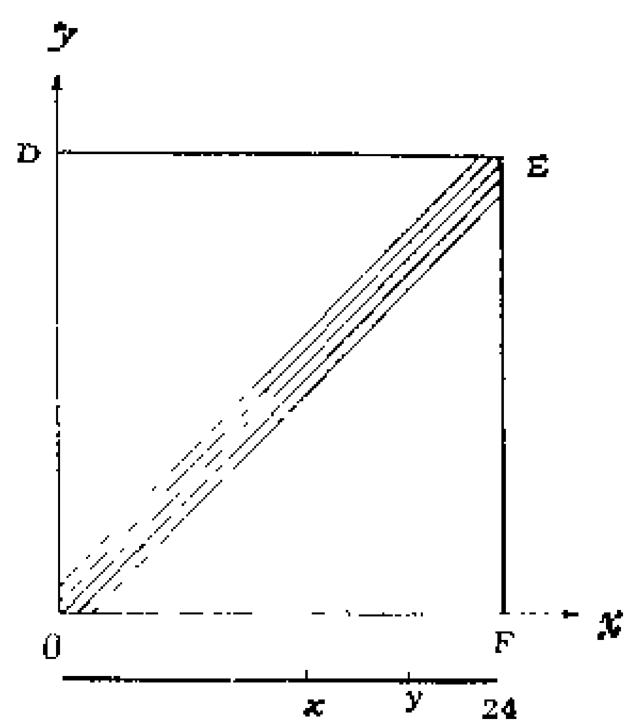
$$l^3-3\cdot\frac{1}{3}\cdot\frac{1}{2}l^3=\frac{1}{2}l^3$$

所以，所求概率为

$$P(A_0)=\frac{\frac{1}{2}l^3}{l^3}=\frac{1}{2}$$



图（1—27题）



图（1—28题）

28　$A^{\triangle}$"一艘轮船需要等待空出码头"，以 x，y 分别表示甲、乙两艘轮船的到达时刻，由于 $0\leqslant x\leqslant 24$，$0\leqslant y\leqslant 24$，所以，甲，乙到达的一切可能时刻对应边长为24的正方形。而事件A发生的充要条件是

$$y-x\leqslant 1 \quad x-y\leqslant 2$$

它对应图（1—28题）中的阴影部分。所以

$$P(A)=\frac{24^2-\left[\frac{1}{2}\cdot 23^2+\frac{1}{2}22^2\right]}{24^2}$$

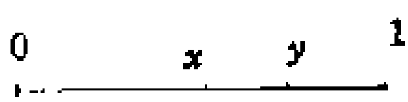
$$\doteq 1-0.87934=0.12066$$

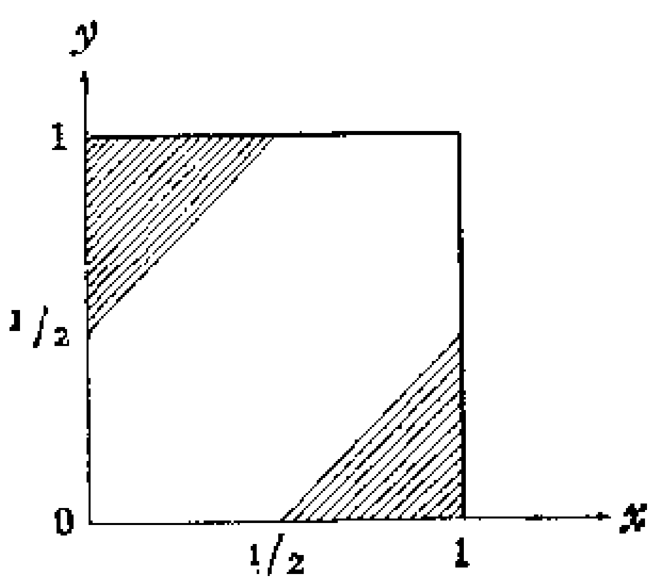
29　$A\triangleq$“一人要等另一个人半小时以上”，令 x，y 分别表示二人的到达时刻，则有

$$0\leqslant x\leqslant 1 \quad 0\leqslant y\leqslant 1$$

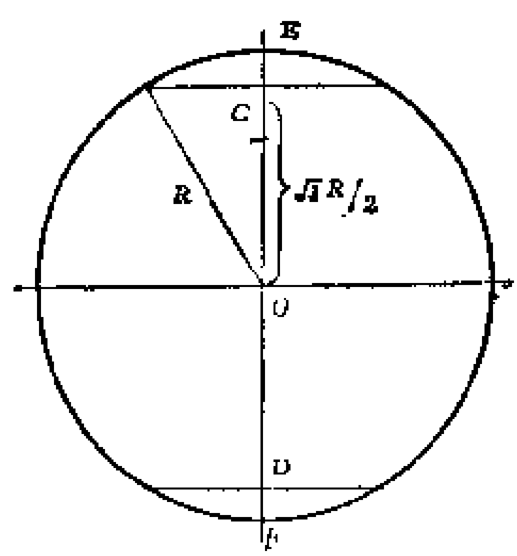
有利于 A 的情况对应图（1—29题）中的阴影部分，所以

$$P(A)=\frac{2\cdot\frac{1}{2}\cdot\left(\frac{1}{2}\right)^2}{1^2}=\left(\frac{1}{2}\right)^2=\frac{1}{4}$$





图（1—29题）



图（1—30题）

30　$A\triangleq$“任意画的弦的长度大于 R”

我们有 $A=$“在直径 EF 上任意投点落在区间 CD 中”

所以

$$P(A)=\frac{CD\text{ 的长}}{EF\text{ 的长}}=\frac{2\sqrt{R^2-\left(\frac{R}{2}\right)^2}}{2R}=\frac{\sqrt{3}}{2}$$

31　$A\triangleq$ "3 个都是白球"

$A_i\triangleq$ "第 i 个是白球" $i=1,2,3$，则　$A=A_1A_2A_3$，由乘法公式知

$$P(A)=P(A_1A_2A_3)=P(A_1)\cdot P(A_2|A_1)\cdot P(A_3|A_1A_2)$$
$$=\frac{a}{a+b}\cdot\frac{a-1}{a+b-1}\cdot\frac{a-2}{a+b-2}$$

此题也可以直接用古典概率的计算公式计算.

$$P(A)=\frac{k}{n}=\frac{{}_aP_3}{{}_{a+b}P_3}=\frac{a(a-1)(a-2)}{(a+b)(a+b-1)(a+b-2)}$$

32　$A\triangleq$ "3 个球中头两个是黑球，第 3 个是白球"

$A_i\triangleq$ "第 i 个球是黑球" $i=1,2,3$，我们有　$A=A_1A_2\overline{A_3}$，由乘法公式知

$$P(A)=P(A_1A_2\overline{A_3})=P(A_1)\cdot P(A_2|A_1)\cdot P(\overline{A_3}|A_1A_2)$$
$$=\frac{b}{a+b}\cdot\frac{b+c}{a+b+c+d}\cdot\frac{a+2d}{a+b+2c+2d}$$

33　$P(A\cap B)=P(B)\cdot P(A|B)=0.4\times0.32=0.128$,

$P(A\cup B)=P(A)+P(B)-P(AB)=0.3+0.4-0.128=0.572$

$$P(\overline{AB})=1-P(AB)=1-0.128=0.872$$
$$P(\overline{A\cup B})=1-P(A\cup B)=1-0.572=0.428$$

34　$A_1\triangleq$ "取第一袋（5 白 7 红）"

$A_2\triangleq$ "取第二袋（4 白 2 红）

$B\triangleq$ "取得白球"

$A_1\cup A_2=U$，$B=BU=B(A_1\cup A_2)=BA_1\cup BA_2$. 所以

$$P(B)=P(BA_1)+P(BA_2)\text{（因为 }A_1\cap A_2=V\text{）}$$

$$= P(A_1)P(B|A_1) + P(A_2)P(B|A_2)$$

$$= \frac{1}{2} \cdot \frac{5}{12} + \frac{1}{2} \cdot \frac{4}{6} = \frac{13}{24}$$

35 $A_1 \triangleq$ “从袋 1 中取白球”

$A_2 \triangleq$ “从袋 1 中取红球”

$B \triangleq$ “从袋 2 中取白球”

则
$$P(B) = P(BA_1) + P(BA_2)$$
$$= P(A_1)P(B|A_1) + P(A_2)P(B|A_2)$$
$$= \frac{x}{x+y} \cdot \frac{u+1}{u+v+1} + \frac{y}{x+y} \cdot \frac{u}{u+v+1}$$
$$= \frac{u(x+y)+x}{(x+y)(u+v+1)}$$

36 $A_1 \triangleq$ “共取 2 次”

$A_2 \triangleq$ “共取 3 次”

$A_3 \triangleq$ “共取 4 次”

则我们有：$A_1 =$ “两次均取坏”，$A_2 =$ “第三次取坏，前两次恰有一次取坏”，$A_3 =$ “最后一次取坏，前三次中恰有一次取坏”。

先考虑A_1，从 4 个零件中有次序地取出 2 个，有 $4 \cdot 3$ 种取法，而第一次取坏有 2 种取法，第二次取坏有 1 种取法，所以

$$P(A_1) = \frac{2 \cdot 1}{4 \cdot 3} = \frac{1}{6}$$

其次考虑A_2，从 4 个零件中有次序地取出 3 个，有 $4 \cdot 3 \cdot 2$ 种取法，而头两次中确定一次取坏有$\binom{2}{1}$种确定法，又在确定的一次中取坏零件有$\binom{2}{1}$种取法，在另一次中取好零件有$\binom{2}{1}$种取法，所以

$$P(A_2) = \frac{\binom{2}{1}\binom{2}{1}\binom{2}{1}}{4 \cdot 3 \cdot 2} = \frac{1}{3}$$

按同样分析可得

$$P(A_3)=\frac{\binom{3}{1}\binom{2}{1}\cdot 2}{4!}=\frac{1}{2}$$

37　$A_1 \triangleq$ “取 2 白 1 黑的坛子”

$A_2 \triangleq$ “取1 0黑的坛子”

$A_3 \triangleq$ “取 3 白 1 黑的坛子”

$B \triangleq$ “取得白球”

由题设知，$P(A_1)=\frac{2}{5}$，$P(A_2)=\frac{1}{5}$，$P(A_3)=\frac{2}{5}$；

$$P(B|A_1)=\frac{2}{3}，P(B|A_2)=0，P(B|A_3)=\frac{3}{4}。$$

注意到 $A_1\cup A_2\cup A_3=U$，且 A_1,A_2,A_3互斥，所以，

$$\begin{aligned}P(B)&=P(BU)=P(B(A_1\cup A_2\cup A_3))\\&=P(BA_1)+P(BA_2)+P(BA_3)\\&=P(A_1)\cdot P(B|A_1)+P(A_2)\cdot P(B|A_2)\\&\quad+P(A_3)P(B|A_3)\\&=\frac{2}{5}\cdot\frac{2}{3}+\frac{1}{5}\cdot 0+\frac{2}{5}\cdot\frac{3}{4}=\frac{34}{60}=\frac{17}{30}\end{aligned}$$

38　已知$\bigcup\limits_{i=1}^{n}A_i=U$，$A_i\cap A_j=\phi$，$i\neq j$

$P(B)>0$，$P(C)>0$

$P(A_i),P(B|A_i),P(C|BA_i)$ 皆已知。

求　$P(C|B)$

解　$$P(C|B)=\frac{P(CB)}{P(B)}$$

$$P(B)=P(BU)=P\left(B\left(\bigcup_{i=1}^{n}A_i\right)\right)=P\left(\bigcup_{i=1}^{n}BA_i\right)$$

$$= \sum_{i=1}^{n} P(BA_i) = \sum_{i=1}^{n} P(A_i) \cdot P(B|A_i)$$

$$P(CB) = P(CBU) = P\left(CB \left(\bigcup_{i=1}^{n} A_i \right)\right) = P\left(\bigcup_{i=1}^{n} CBA_i \right)$$

$$= \sum_{i=1}^{n} P(CBA_i) = \sum_{i=1}^{n} P(BA_i) \cdot P(C|BA_i)$$

$$= \sum_{i=1}^{n} P(A_i) \cdot P(B|A_i) \cdot P(C|BA_i)$$

所以，

$$P(C|B) = \frac{\sum_{i=1}^{n} P(A_i) \cdot P(B|A_i) \cdot P(C|BA_i)}{\sum_{i=1}^{n} P(A_i) \cdot P(B|A_i)}$$

39　已知　$P(A) = 0.5, P(B|A) = P(\overline{B}|\overline{A}) = 0.9$，求　$P(A|B)$，$P(A|B)$。

解　$P(B) = P(BA) + P(B\overline{A})$

$$= P(A) \cdot P(B|A) + P(\overline{A}) \cdot P(B|\overline{A})$$

$$= 0.5 \times 0.9 + (1 - 0.5) \times 0.1 = 0.5$$

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(A) \cdot P(B|A)}{P(B)} = \frac{0.5 \cdot 0.9}{0.5} = 0.9$$

$$P(A|\overline{B}) = \frac{P(A\overline{B})}{P(\overline{B})} = \frac{P(A - AB)}{1 - P(B)} = \frac{P(A) - P(AB)}{1 - P(B)}$$

$$= \frac{0.5 - 0.5 \times 0.9}{0.5} = 0.1$$

40　$A_1 \triangleq$ “从1号袋取球”，

$A_2 \triangleq$ “从不是1号袋取球”

$B \triangleq$ “取白球”。

求“这个白球是来自1号袋子”的概率，就是求条件概率

$P(A_1|B)$，由 Bayes 公式有

$$P(A_1|B)=\frac{P(A_1)\cdot P(B|A_1)}{P(A_1)\cdot P(B|A_1)+P(A_2)\cdot P(B|A_2)}$$

$$=\frac{\frac{1}{5}\cdot\frac{1}{3}}{\frac{1}{5}\cdot\frac{1}{3}+\frac{4}{5}\cdot\frac{1}{4}}=\frac{1}{4}$$

41 $$P(B|A)=\frac{P(B)\cdot P(A|B)}{P(A)}$$

式中，$P(B)=0.0004$，$P(A|B)=0.95$，欲算出 $P(A)$，用全概率公式得

$$P(A)=P(B)\cdot P(A|B)+P(\overline{B})\cdot P(A|\overline{B})$$

$$P(B)\cdot P(A|B)+[1-P(B)]\cdot[1-P(\overline{A}|\overline{B})]$$

$$=0.0004\times 0.95+0.9996\times 0.1=0.10034,$$

所以，

$$P(B|A)=\frac{0.00038}{0.10034}=0.00379$$

42 $A_1\triangleq$“取的枪试射过”，

$A_2\triangleq$“取的枪末试射过”，

$B\triangleq$“射击命中”，

$$P(B)=P(A_1)\cdot P(B|A_1)+P(A_2)\cdot P(B|A_2)$$

$$=\frac{2}{9}\cdot\frac{8}{10}+\frac{7}{9}\cdot\frac{1}{10}=\frac{23}{90}$$

$$P(A_1|B)=\frac{P(A_1)\cdot P(B|A_1)}{P(B)}=\frac{\frac{16}{90}}{\frac{23}{90}}=\frac{16}{23}$$

43 $A_1\triangleq$“取出正品” $P(A_1)=\frac{10}{100}$

$A_2 \triangleq$ "取出非正品"　　$P(A_2)=\frac{90}{100}$

$B \triangleq$ "使用 n 次均无故障"

按题要求应　$P(A_1|B)\geqslant 0.70$，而

$$P(A_1|B)=\frac{P(A_1)P(B|A_1)}{P(A_1)P(B|A_1)+P(A_2)P(B|A_2)}$$

$$=\frac{0.1\times 1}{0.1\times 1+0.9\times(0.9)^n}$$

所以应是　$\frac{0.1}{0.1+(0.9)^{n+1}}\geqslant 0.70$，$0.043\geqslant(0.9)^{n+1}$

得

$$n\geqslant 29$$

44　$A_1 \triangleq$ "第一次取出黑球"，

$A_2 \triangleq$ "第一次取出红球"，

$B \triangleq$ "第二次取出黑球"．

$$P(B)=P(A_1)\cdot P(B|A_1)+P(A_2)\cdot P(B|A_2)$$

$$=\frac{b}{b+r}\cdot\frac{b+c}{b+r+c}+\frac{r}{b+r}\cdot\frac{b}{b+r+c}=\frac{b}{b+r}$$

$$P(A_1|B)=\frac{P(A_1B)}{P(B)}=\frac{\frac{b}{b+r}\cdot\frac{b+c}{b+r+c}}{\frac{b}{b+r}}$$

$$=\frac{b+c}{b+r+c}$$

45　$A \triangleq$ "前面 n_1 次取得黑球，后面 $n_2=n-n_1$ 次取得红球"

$A_i \triangleq$ "第 i 次取得黑球"　$i=1,2,\cdots,n$，

我们有

$$A = A_1A_2\cdots A_{n_1}\overline{A}_{n_1+1}\overline{A}_{n_1+2}\cdots\overline{A}_n$$

但 $$P(A_1) = \frac{b}{b+r},\quad P(A_2|A_1) = \frac{b+c}{b+r+c},\cdots,$$

$$P\{A_{n_1}|A_1A_2\cdots A_{n_1-1}\} = \frac{b+(n_1-1)c}{b+r+(n_1-1)c},$$

$$P\{\overline{A}_{n_1+1}|A_1A_2\cdots A_{n_1}\} = \frac{r}{b+r+n_1c},\cdots,$$

$$P\{\overline{A}_n|A_1\cdots A_{n_1}\overline{A}_{n_1+1}\cdots\overline{A}_{n-1}\} = \frac{r+(n_2-1)c}{b+r+(n-1)c}$$

由乘法公式可知

$$P\{A\} = \frac{b}{b+r}\cdot\frac{b+c}{b+r+c}\cdots\frac{b+(n_1-1)c}{b+r+(n_1-1)c}$$

$$\cdot\frac{r}{b+r+n_1c}\cdots\frac{r+(n_2-1)c}{b+r+(n-1)c}$$

46 用数学归纳法．

$n=1$时，$P(A_1) = \dfrac{b}{b+r}$，命题真．

假定命题对 $n=k$ 时成立，即第 k 次取得黑球的概率 $P\{A_k\} = \dfrac{b}{b+r}$，往证命题当 $n=k+1$ 时成立．

$$P(A_{k+1}) = P(A_1A_{k+1}) + P(\overline{A}_1A_{k+1})$$

$$= P(A_1)\cdot P(A_{k+1}|A_1) + P(\overline{A_1})\cdot P(A_{k+1}|\overline{A_1})$$

$P\{A_{k+1}|A_1\}$为在第一次取出黑球的条件下，第 $k+1$次取出黑球的条件概率，它可以看成是从有 $b+c$ 个黑球，r 个红球的坛子里经过k 次取球，取出黑球的概率．由归纳假定，它等于$\dfrac{b+c}{b+r+c}$．

同理，$P(A_{k+1}|\overline{A_1}) = \dfrac{b}{b+r+c}$，所以

$$P(A_{k+1}) = \frac{b}{b+r}\cdot\frac{b+c}{b+r+c} + \frac{r}{b+r}\cdot\frac{b}{b+r+c}$$

$$= \frac{b}{b+r}$$

由归纳法原理知，命题得证。

47 $$P(A)=P(C)\cdot P(A|C)+P(\overline{C})\cdot P(A|\overline{C})$$
$$=\frac{1}{2}(0.9+0.2)=0.55$$

$$P(B)=P(C)\cdot P(B|C)+P(\overline{C})\cdot P(B|\overline{C})$$
$$=\frac{1}{2}(0.9+0.1)=0.5$$

$$P(A\cap B)=P(C)\cdot P(AB|C)+P(\overline{C})\cdot P(AB|\overline{C})$$
$$=P(C)\cdot P(A|C)\cdot P(B|C)+P(\overline{C})\cdot P(A|\overline{C})\cdot P(B|\overline{C})$$
（由 A,B 条件独立）
$$=\frac{1}{2}[0.9^2+0.2\times 0.1]=0.415$$

由于 $P(A\cap B)=0.415\neq 0.55\times 0.5=P(A)\cdot P(B)$，所以 A，B 不独立。

48 $A\triangleq$ “不出废品”

$A_i\triangleq$ “第 i 道工序不出废品” $i=1,2,3$.

我们有，$A=A_1A_2A_3$。已假定诸A_i相互独立，由乘法公式可知
$$P(A)=P(A_1)P(A_2)P(A_3)$$
$$=[1-P(\overline{A_1})][1-P(\overline{A_2})][1-P(\overline{A_3})]$$
$$=[1-0.02]^3=0.98^3=0.94112$$

49 $$P(A\cup B)=P(A)+P(B)-P(AB)$$
$$=P(A)+P(B)-P(A)P(B)\quad（由 A，B 独立）$$
$$=P(A)+P(B)[1-P(A)]$$

所以，
$$P(B)=\frac{P(A\cup B)-P(A)}{1-P(A)}=\frac{0.6-0.4}{1-0.4}=\frac{0.2}{0.6}=\frac{1}{3}$$

50 $A\triangleq$ “二人获得相同数的正面”

$A_i \triangleq$ "二人各得 i 个正面"，$i=0,1,2,3$

我们有，

$$A=A_0\cup A_1\cup A_2\cup A_3$$

所以

$$P(A)=P(A_0)+P(A_1)+P(A_2)+P(A_3)$$

$$=\frac{1}{8}\cdot\frac{1}{8}+\frac{3}{8}\cdot\frac{3}{8}+\frac{3}{8}\cdot\frac{3}{8}+\frac{1}{8}\cdot\frac{1}{8}=\frac{20}{64}$$

51　$A\triangleq$ "一个二进数是错误的"

$A_i\triangleq$ "第 i 位数码是错误的"，$i=1,2,\cdots,n$

我们有，

$$A=A_1\cup A_2\cup\cdots\cup A_n$$

$$P(A)=1-P\{\overline{A}\}=1-P\{\overline{A}_1\cap\overline{A}_2\cap\cdots\cap\overline{A}_n\}$$

$$=1-\prod_{i=1}^{n}P(\overline{A}_i)=1-(1-p)^n\qquad(\text{由诸 }\overline{A}_i\text{ 独立})。$$

52　$A\triangleq$ "二阶行列式的值大于 0"

若记二阶行列式为，

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

则，$A=(a=1,\ d=1,\ b,\ c$ 中至少有一个是 0）

由乘法公式知，

$$P(A)=P(a=1)\cdot P(d=1)\cdot P(b,c\text{ 中至少有一个是}0)$$

$$=\frac{1}{2}\cdot\frac{1}{2}\cdot\left[1-\frac{1}{2}\cdot\frac{1}{2}\right]=\frac{1}{4}\cdot\frac{3}{4}=\frac{3}{16}$$

53　$A\triangleq$ "电流由 L 流到 R"

$A_i\triangleq$ "开关 A_i 闭上"，$i=1,2,3,4,5$

我们有，

$$A = A_5 \cup A_1A_2 \cup A_3A_4$$

$$\Rightarrow \overline{A} = \overline{A_5} \cap \overline{A_1A_2} \cap \overline{A_3A_4}$$

$$\Rightarrow P(\overline{A}) = P(\overline{A_5}) \cdot P(\overline{A_1A_2}) \cdot P(\overline{A_3A_4})$$

$$= [1 - P(A_5)] \cdot [1 - P(A_1A_2)] \cdot [1 - P(A_3A_4)]$$

$$= [1 - P(A_5)][1 - P(A_1) \cdot P(A_2)]$$

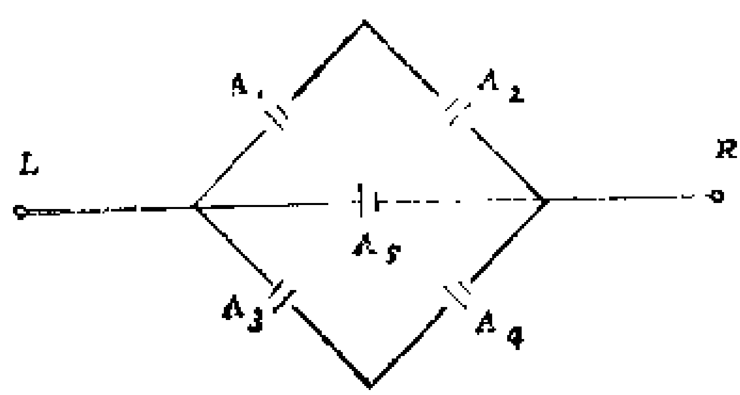
$$[1 - P(A_3) \cdot P(A_4)]$$

$$= (1 - p)(1 - p^2)(1 - p^2) = (1 - p)(1 - p^2)^2$$

所以，

$$P(A) = 1 - P(\overline{A}) = 1 - (1 - p)(1 - p^2)^2$$

$$= p + 2p^2 - 2p^3 - p^4 + p^5$$



图（1—53题）

54　$A \triangleq$ "当25支步枪同时开火时，击落敌机"

$A_i \triangleq$ "第 i 号步枪开火击落敌机"，$i = 1, 2, \cdots, 25$

我们有，

$$A = A_1 \cup A_2 \cup A_3 \cup \cdots \cup A_{25}$$

于是，

$$P(A) = 1 - P(\overline{A}) = 1 - P(\overline{A_1} \cap \overline{A_2} \cap \cdots \cap \overline{A_{25}})$$

$$= 1 - \prod_{i=1}^{25} P(\overline{A_i}) = 1 - [1 - 0.008]^{25} = 1 - 0.992^{25}$$

$$=1-0.81807=0.18193$$

55　$A\triangleq$ "n门炮至少有一门击中目标"

$A_i\triangleq$ "第 i 号炮击中目标"，$i=1,2,\cdots,n$

我们有，

$$A=A_1\cup A_2\cup\cdots\cup A_n$$

于是，

$$P(A)=1-P(\overline{A})=1-P(\overline{A_1}\cap\overline{A_2}\cap\cdots\cap\overline{A_n})$$

$$=1-\prod_{i=1}^{n}P(\overline{A_i})=1-[1-0.6]^n=1-0.4^n$$

按题意，要求 $P(A)\geqslant 0.99$，即 $1-(0.4)^n\geqslant 0.99\Rightarrow(0.4)^n\leqslant 0.01$ $\Rightarrow n\geqslant 6$。

56　$W_1\triangleq$ "第一种设计正常工作"

$W_2\triangleq$ "第二种设计正常工作"

框图中第一行的四个元件（由左到右）正常工作分别记为A_1，A_2，A_3，A_4；第二行的四个元件正常工作分别记为 B_1，B_2，B_3，B_4。

由题设有，

$$P(\overline{A_1})=P(\overline{B_1})=0.05,\ P(\overline{A_2})=P(\overline{B_2})=0.01,$$

$$P(\overline{A_3})=P(\overline{B_3})=0.10,\ P(\overline{A_4})=P(\overline{B_4})=0.04.$$

第一种设计正常工作的概率为

$$P(W_1)=P(A_1A_2A_3A_4\cup B_1B_2B_3B_4)$$

$$=1-P\ (\overline{A_1A_2A_3A_4}\cap\overline{B_1B_2B_3B_4})$$

$$=1-[1-P(A_1A_2A_3A_4)]\cdot[1-P(B_1B_2B_3B_4)]$$

$$=1-\left[1-\prod_{i=1}^{4}P(A_i)\right]\cdot\left[1-\prod_{i=1}^{4}P(B_i)\right]$$

$$=1-[1-0.95\times 0.99\times 0.90\times 0.96]$$

$$\cdot[1-0.95\times 0.99\times 0.90\times 0.96]$$

$$=0.96488$$

第二种设计正常工作的概率为

$$P(W_2)=P((A_1\cup B_1)(A_2\cup B_2)(A_3\cup B_3)(A_4\cup B_4))$$

$$=\prod_{i=1}^{4}P(A_i\cup B_i)=\prod_{i=1}^{4}[1-P(\overline{A_i\cup B_i})]$$

$$=\prod_{i=1}^{4}[1-P(\overline{A_i}\cap\overline{B_i})]$$

$$=\prod_{i=1}^{4}[1-P(\overline{A_i})\cdot P(\overline{B_i})]$$

$$=[1-0.05^2][1-0.01^2][1-0.1^2][1-0.04^2]$$

$$=0.9975\times0.9999\times0.99\times0.9984=0.98585$$

由于 $P(W_2)>P(W_1)$ 所以，第二种设计方案更可靠。

第二章 离散型随机变量

从第二章开始，我们进入对随机变量的研究。引进随机变量的概念以后，我们对随机现象的研究就更加突出了数量这一侧面。以后会看到，对随机事件的研究完全可以转化为对随机变量的研究。

§2.1 随机变量与概率分布

(一) 随机变量

我们知道，进行某种试验，它可能有各种不同的结果。一般来说，试验的每一个基本事件都对应一个数。比如，在一批产品中随机取10件，考虑其中的废品数，这时，基本事件是

$$\{\omega_i\}=\text{“有 } i \text{ 件废品”},\ i=0,\ 1,\ 2,\ \cdots,\ 10$$

它们分别对应数值 0，1，2，…，10，这些数就是废品数的取值范围。由于废品数可能取 0，1，2，…，10中的任一个值，所以它是一个变量，又由于试验结果具有偶然性，所以这个变量取哪一个值也具有偶然性。象废品数这样的，随着试验结果的不同以偶然方式取值的量叫做随机变量。显然随机变量是定义在基本事件空间上的函数。

我们可以举出很多随机变量的例子。比如，在n次射击中命中靶子的次数；电话交换台在一段时间内接到的呼唤数；某种另件的尺寸；同一品种的小麦在条件相同的土地上的亩产量；一条河流每年最大流量，等等，都是随机变量。随机变量用希腊字母ξ，η，ζ，…表示。

有些试验的结果，看来并不具有数量涵义，比如，扔一枚硬币，有两个基本事件："出正面"和"出反面"。我们也可使它的结果与数对应，比如可以令"出正面"对应1，"出反面"对应0。类似地，在射击中"命中"，"不命中"；在抽检一件产品时，"是好品"，"是次品"；在观察某一天的天气时，"下雨"，"不下雨"，等等，都可按同样方法使其结果与数对应。因此，对这类试验也可以在基本事件空间上定义随机变量。一般地，若试验只有两个结果A与$\overline{A}$，这试验叫做贝努利（*Bernoulli*）试验，对贝努利试验，函数

$$\chi_A=\begin{cases}1 & \text{若}A\text{发生}\\0 & \text{若}A\text{不发生}\end{cases}$$

就是与其对应的随机变量。把它叫做事件A的示性函数。

对于一个随机变量，如果我们仅仅知道它可能取什么值是不够的，更有意义的是应当知道它取值在某一范围里的可能性多大，比如，$a\leqslant\xi\leqslant b$的可能性多大。我们规定$a\leqslant\xi\leqslant b$的可能性就是使$a\leqslant\xi\leqslant b$成立的那些基本事件集合的概率。比如，若ξ表示上面讲过的10件产品中的废品数，则使$3\leqslant\xi\leqslant 5$成立的基本事件是ω_3，ω_4，ω_5，所以$P(3\leqslant\xi\leqslant5)=P\{\omega_3,\ \omega_4,\ \omega_5\}$。

由于概率只是对事件域$\mathscr{F}$中的事件定义的，所以要使$a\leqslant\xi\leqslant b$有确定的概率，就必须使$a\leqslant\xi\leqslant b$所对应的基本事件集合属于$\mathscr{F}$。因此，给出随机变量的确切定义如下。

定义2.1.1　设$(\Omega,\mathscr{F},P)$是一概率空间，$\xi(\omega)$是定义在Ω上的单值实函数，如果对于任一实数x有

$$(\omega:\xi(\omega)\leqslant x)\in\mathscr{F}$$

则$\xi(\omega)$或简写为ξ叫做随机变量。

今后为叙述简便，和事件$(\omega:\xi(\omega)\leqslant x)$一样，把$\xi\leqslant x$也叫做事件。

在本节一开始，我们举了一些随机变量的例子。其中有的随机变量所能取值全体是有限个（如废品数），有的随机变量所能取值

全体是可列个（如交换台接到的呼唤数），这两种随机变量统称为离散型随机变量．象另件的尺寸这个随机变量，它所能取值连续地充满某一区间，我们称它为连续型随机变量．经常遇到的随机变量主要就是离散型与连续型两种．这一章讲离散型随机变量．下一章讲连续型随机变量．

（二）概率分布

对于一个离散型随机变量，要掌握它的变化规律，首先要了解它可能取什么值，除此之外，更主要的，还要了解它取各可能值的概率是多少．

如果一个离散型随机变量的可能取值为x_1，x_2，…，x_n，…，而且“$\xi=x_k$”的概率为

$$P(\xi=x_k)=p(x_k),\quad k=1,\ 2,\ \cdots \tag{2.1.1}$$

则（2.1.1）不仅告诉我们ξ能取什么值，而且还指出了它以多大的概率取这些值，所以它把随机变量ξ完整地描述出来了．我们称（2.1.1）为随机变量ξ的概率函数．有时将（2.1.1）写成下面形式

$$\begin{pmatrix} x_1, & x_2, & \cdots & x_n, & \cdots \\ p(x_1), & p(x_2), \cdots, & & p(x_n), \cdots \end{pmatrix} \tag{2.1.2}$$

（2.1.2）称为ξ的分布列，（2.1.2）和（2.1.1）都是ξ的概率分布（或简称分布）

若一个随机变量ξ的概率分布已知，则ξ取值在任一区间的概率都可以知道．比如若随机变量ξ的概率分布为

$$\begin{bmatrix} 0, & 1, & 2, & 3, & 4 \\ \frac{1}{16}, & \frac{4}{16}, & \frac{6}{16}, & \frac{4}{16}, & \frac{1}{16} \end{bmatrix}$$

则

$$P(\xi\leqslant 2)=\sum_{k=0}^{2} P(\xi=k)=11/16$$

$$P(1\leqslant\xi\leqslant 3)=\sum_{k=1}^{3} P(\xi=k)=14/16=7/8$$

$$P(0.6\leqslant\xi\leqslant 3.1)=\sum_{k=1}^{3}P(\xi=k)=7/8$$

一般地，若已知ξ的概率分布为（2.1.2），则对任意实数$a<b$，由于$(a\leqslant\xi\leqslant b)=\bigcup\limits_{a\leqslant x_i\leqslant b}(\xi=x_i)$，且$(\xi=x_i)$互斥，所以有

$$P(a\leqslant\xi\leqslant b)=\sum_{a\leqslant x_i\leqslant b}P(\xi=x_i)$$

等式右端是对一切满足$a\leqslant x_i\leqslant b$的$x_i$求和。由此可知，$\xi$取值在任一区间的概率都可由其概率分布求出。所以，我们说概率分布完全描述了随机变量的变化规律。

由概率的基本性质可知，对于任一概率分布都有下面两个性质。

（1）$p(x_k)\geqslant 0$，$k=1, 2, 3, \cdots$

（2）$\sum\limits_{k=1}^{\infty}p(x_k)=1$

这是由于事件$(\xi=x_1)$，$(\xi=x_2)$，…互斥，且其和为必然事件，所以

$$\sum_{k=1}^{\infty}p(x_k)=P\left(\bigcup_{k=1}^{\infty}(\xi=x_k)\right)=P(U)=1$$

下面举几个概率分布的例子。

例2.1.1（单点分布）　常数c可以看做是最简单的特殊的随机变量，即在每次试验中必然取值c的随机变量。它的分布列是

$$\begin{pmatrix} c \\ 1 \end{pmatrix}$$

例2.1.2（两点分布或贝努利分布）　可能取值只有两个（0和1）的随机变量叫做服从两点分布的随机变量。若$P(\xi=1)=p$，$P(\xi=0)=1-p=q$，$0\leqslant p\leqslant 1$，则其分布列是

$$\begin{pmatrix} 0, & 1 \\ q, & p \end{pmatrix}$$

例2.1.3　（超几何分布）一批产品共N件，其中有M件是废品，从中任取n件$(n\leqslant N)$，以ξ表示n件中的废品数，则由§1.3

的例1.3.2知道，ξ的概率分布是

$$P(\xi=m)=\frac{\binom{M}{m}\binom{N-M}{n-m}}{\binom{N}{n}}\qquad \begin{matrix} m\leqslant n \\ m\leqslant M \end{matrix}\qquad n-m\leqslant N-M$$

这个分布叫超几何分布。

例2.1.4（几何分布） 独立地重复 $P(A)=p$，$P(\overline{A})=1-p=q$ 的贝努利试验，考虑直到第一次出现A所需要的试验次数ξ。ξ的可能取值为1，2，…，$(\xi=k)$ 表示前$k-1$次试验都没有出现A，而第k次试验才出现A，显然它的概率为

$$P(\xi=k)=q^{k-1}p,\qquad k=1,\ 2,\ \cdots$$

这个分布叫做几何分布。

（三）联合分布与独立性概念

有时，对应一个试验有两个或两个以上的随机变量。比如，检查圆筒形零件时，如果是考察它的直径和长度，则这种检查结果就对应两个随机变量。如果对应一个试验有n个随机变量ξ_1，ξ_2，…，ξ_n，则我们称它们的总体（ξ_1，ξ_2，…，ξ_n）为n维随机变量（或称n维随机向量）。下面我们着重讨论二维随机变量。

定义2.1.2 若ξ，η为定义在同一概率空间$(\Omega,\ \mathscr{F},\ P)$上的两个随机变量，则称$(\xi,\ \eta)$为二维随机变量。

设$(\xi,\ \eta)$为二维随机变量，ξ，η的可能取值分别为x_1，x_2，…，y_1，y_2，…，其对应的概率分布分别为$g(x_i)=P(\xi=x_i)$，$h(y_k)=P(\eta=y_k)$，我们用$P(x_i,\ y_k)$表示$(\xi=x_i)$与$(\eta=y_k)$同时发生的概率，即

$$p(x_i,\ y_k)=P\{(\xi=x_i)\cap(\eta=y_k)\}\qquad(2.1.3)$$

则函数（2.1.3）叫做二维随机变量$(\xi,\ \eta)$的联合分布。

现在把事件间独立的概念推广到随机变量上去。

定义2.1.3 若对随机变量ξ与η的所有可能取值x_i与y_k，事件$(\xi=x_i)$与$(\eta=y_k)$都是独立的，即

$$P\{(\xi=x_i)\cap(\eta=y_k)\}=P(\xi=x_i)\cdot P(\eta=y_k)$$

则称 ξ 与 η 相互独立。

下面研究二维随机变量 $(\xi,\ \eta)$ 的联合分布 $p(x_i,\ y_k)$ 与 ξ 的分布 $g(x_i)$，η 的分布 $h(y_k)$ 之间的关系。

首先，由 $p(x_i,\ y_k)$ 可以决定 $g(x_i)$，$\eta(y_k)$，这是因为

$$(\eta=y_k)\quad k=1,\ 2,\ \cdots \text{互斥，且} \bigcup_k(\eta=y_k)=U$$

所以

$$g(x_i)=P\ (\xi=x_i)=P\{\sum_k(\xi=x_i)(\eta=y_k)\}$$

$$=\sum_k P\{(\xi=x_k)(\eta=y_k)\}=\sum_k p(x_i,y_k)$$

同理 (2.1.4)

$$h(y_k)=\sum_i p(x_i,\ y_k) \tag{2.1.5}$$

由 (2.1.4)，(2.1.5) 可知，二维随机变量 $(\xi,\ \eta)$ 的联合分布能决定 ξ 与 η 各自的分布。

反之，由 $g(x_i)$ 与 $h(y_k)$ 能否决定 $p(x_i,\ y_k)$ 呢？

如果 ξ 与 η 独立，则由 ξ 与 η 的分布 $g(x_i)$ 与 $h(y_k)$ 能够决定 $(\xi,\ \eta)$ 的联合分布 $p(x_i,\ y_k)$，这是因为

$$p(x_i,\ y_k)=P\{(\xi=x_i)\cap(\eta=y_k)\}$$
$$=P(\xi=x_i)\cdot P(\eta=y_k)=g(x_i)\cdot h(y_k) \tag{2.1.6}$$

但在一般情况下，由 $g(x_i)$ 和 $h(y_k)$ 决定不了 $p(x_i,\ y_k)$，这是因为联合分布 $p(x_i,\ y_k)$ 除了反映 ξ，η 各自的变化规律外，还反映 ξ 与 η 之间的相互关系，它反映的是 $(\xi,\ \eta)$ 作为一个整体的变化规律。总之，二维随机变量 $(\xi,\ \eta)$ 的联合分布决定 ξ 与 η 各自的分布，但是，只有当 ξ 与 η 相互独立时，ξ 与 η 的各自分布才能决定 $(\xi,\ \eta)$ 的联合分布。有时要同时讨论 $(\xi,\ \eta)$ 的联合分布与 ξ，η 的各自分布，这时为了说起来方便，把 ξ 与 η 各自的分布叫做边缘分布。把 (2.1.4) 和 (2.1.5) 表示成下表看起来更直观一些。

联合分布		ξ				η的分布
		x_1	x_2	$\cdots$	$x_n\cdots$	
η	y_1	$p(x_1,y_1)$	$p(x_2,y_1)$	$\cdots$	$p(x_n,y_1)\cdots$	$h(y_1)=\sum\limits_j p(x_j,y_1)$
	y_2	$p(x_1,y_2)$	$p(x_2,y_2)$	$\cdots$	$p(x_n,y_2)\cdots$	$h(y_2)=\sum\limits_j p(x_j,y_2)$
	$\vdots$	$\cdots\cdots$	$\cdots\cdots$	$\cdots\cdots$	$\cdots\cdots$	$\vdots$
	y_n	$p(x_1,y_n)$	$p(x_2,y_n)$	$\cdots$	$p(x_n,y_n)\cdots$	$h(y_n)=\sum\limits_j p(x_j,y_n)$
	$\vdots$	$\cdots\cdots$	$\cdots\cdots$	$\cdots\cdots$	$\cdots\cdots$	$\vdots$
ξ的分布		$g(x_1)=\sum\limits_k P(x_1,y_k)$	$g(x_2)=\sum\limits_k p(x_2,y_k)$	$\cdots$	$g(x_n)=\sum\limits_k P(x_n,y_k)$	

在上表中，右边一列是把每一行的 $p(x_j, y_k)$按各x_j相加得到的，它就是η的分布列 $h(y_k)$。同样，在表中下面一行是把每一列的 $p(x_j, y_k)$按各 y_k 相加得到的，它就是ξ的分布列$g(x_j)$。

例2.1.5 若ξ_1，ξ_2相互独立且服从相同的两点分布$\begin{pmatrix} 0, & 1 \\ q, & p \end{pmatrix}$，求$(\xi_1, \xi_2)$ 的联合分布。

解 二维随机变量(ξ_1, ξ_2)的所有可能取值是(0，0)，(0，1)，(1，0) 与 (1，1) 。由独立性知道它的联合分布为

$$P\{(\xi_1=0)\cdot(\xi_2=0)\}=P(\xi_1=0)\cdot P(\xi_2=0)=q\cdot q=q^2;$$

同样得 $P\{(\xi_1=0)\cdot(\xi_2=1)\}=qp;$

$$P\{(\xi_1=1)\cdot(\xi_2=0)\}=qp;$$

$$P\{(\xi_1=1)\cdot(\xi_2=1)\}=p^2$$

可以通过这个实例验证 (2.1.4) 和 (2.1.5) 。

例2.1.6 把标号为1，2，3的三个球随机地放入三个盒中。令ξ为第一个盒中的球数，η为已放有球的盒子数。求(ξ, η)的联合分布及ξ，η的边缘分布。

解　ξ可能取值为0，1，2，3，η可能取值为1，2，3。现在求

$P(\xi=i,\ \eta=j)$，$i=0,\ 1,\ 2,\ 3$，$j=1,\ 2,\ 3$。

先求$P(\xi=0,\ \eta=1)$。由于事件$(\xi=0,\ \eta=1)=$（第一个盒中没有球，而其余两盒只有一个盒有球）包含两种情况，即第二个盒中有三个球或第三个盒中有三个球，所以 $P(\xi=0,\ \eta=1)=\frac{2}{27}$（因为三个球放入三个盒中总共有 $3^3=27$ 种情况）。

再求$P(\xi=0,\ \eta=2)$。由于事件$(\xi=0,\ \eta=2)=${第一个盒中没有球，而其余两盒中都有球）包含 6 种情况，即第二个盒中有两个球而第三个盒中有一个球或第二个盒中有一个球而第三个盒中有两个球。所以$P(\xi=0,\ \eta=2)=\frac{6}{27}$

……

可将所求结果列成如下的表

		ξ 0	1	2	3	η的分布
	1	$\frac{2}{27}$	0	0	$\frac{1}{27}$	$\frac{1}{9}$
η	2	$\frac{6}{27}$	$\frac{6}{27}$	$\frac{6}{27}$	0	$\frac{2}{3}$
	3	0	$\frac{6}{27}$	0	0	$\frac{2}{9}$
ξ的分布		$\frac{8}{27}$	$\frac{12}{27}$	$\frac{6}{27}$	$\frac{1}{27}$	

ξ的分布为

$$\begin{pmatrix} 0, & 1, & 2, & 3 \\ \frac{8}{27}, & \frac{12}{27}, & \frac{6}{27}, & \frac{1}{27} \end{pmatrix}$$

η的分布为

$$\begin{pmatrix} 1, & 2, & 3 \\ \frac{1}{9}, & \frac{6}{9}, & \frac{2}{9} \end{pmatrix}$$

（四）条件分布

对于两个随机事件我们曾定义了它们的条件概率，对应地，对于两个随机变量也可定义它们的条件分布。

已知二维随机变量(ξ, η)的联合分布为$p(x_i, y_k)=P(\xi=x_i, \eta=y_k)$，则在已知$\xi=x_i$的条件下$(g(x_i)=P(\xi=x_i)>0)$，事件$(\eta=y_k)$的条件概率为

$$P(\eta=y_k|\xi=x_i)=\frac{P(\xi=x_i,\ \eta=y_k)}{P(\xi=x_i)}$$

$$=\frac{p(x_i,\ y_k)}{\sum_{k=1}^{\infty}p(x_i, y_k)} \qquad (2.1.7)$$

式（2.1.7）定义了在$\xi=x_i$条件下η的条件分布。在一般情况下，它不同于$h(y_k)=P(\eta=y_k)$。

由上式可知联合分布一般可表示为

$$P(\xi=x_i, \eta=y_k)=P(\xi=x_i)P(\eta=y_k|\xi=x_i) \qquad (2.1.8)$$

或写成

$$p(x_i,\ y_k)=g(x_i)\cdot h(y_k|x_i)$$

当两个随机变量ξ，η相互独立时，已知有（2.1.6），因此，这时有

$$h(y_k|x_i)=h(y_k)$$

容易验证，条件分布与分布同样具有如下两个性质。

（1）$P(\eta=y_k|\xi=x_i)\geqslant 0$，$k=1, 2, \cdots$

（2）$\sum_{k=1}^{\infty}P(\eta=y_k|\xi=x_i)=1$

对于分布具有的其它性质，条件分布也有类似的性质，这里就不一一细述了。

例2.1.7　已知随机变量ξ的概率分布为

$$P(\xi=j)=e^{-\lambda}\frac{\lambda^j}{j!},\ j=0,\ 1,\ 2,\ \cdots;\ \lambda>0$$

而η的条件概率分布为

$$P(\eta=k|\xi=j)=\binom{j}{k}p^k(1-p)^{j-k}$$

$$k=0,\ 1,\ \cdots,\ j,\ 0<p<1$$

求η的概率分布。

解　先求$(\xi,\ \eta)$的联合分布$P(\xi=j,\ \eta=k)$

$$P(\xi=j,\ \eta=k)=P(\xi=j)\cdot P(\eta=k\,|\,\xi=j)$$

$$=e^{-\lambda}\frac{\lambda^j}{j!}\cdot\binom{j}{k}p^k(1-p)^{j-k}$$

于是

$$P(\eta=k)=\sum_{j=k}^{\infty}e^{-\lambda}\frac{\lambda^j}{j!}\cdot\frac{j!}{k!(j-k)!}p^k(1-p)^{j-k}$$

$$=e^{-\lambda}\frac{(\lambda p)^k}{k!}\sum_{j=k}^{\infty}\frac{[\lambda(1-p)]^{j-k}}{(j-k)!}$$

$$=e^{-\lambda}\frac{(\lambda p)^k}{k!}e^{\lambda(1-p)}$$

$$=e^{-\lambda p}\frac{(\lambda p)^k}{k!},\quad k=0,\ 1,\ 2,\ \cdots$$

在离散型随机变量的分布中有两种分布是最常见也是最重要的，即二项分布和普阿松分布，下面两节将分别讨论这两种分布。

§2.2　二项分布

二项分布来源于n重贝努利试验，所谓n重贝努利试验，就是将贝努利试验独立地重复n次所构成的试验。在n重贝努利试验中，事件A（为贝努利试验的一个结果）发生的次数是一个随机变量，现在来考虑它的分布。先看一个具体例子。

独立地进行5次射击，设ξ为在5次射击中命中靶子的次数，假定已知在每次射击中"命中靶子"的概率都是0.9，求ξ的概率分布。

ξ可能取值是0，1，2，3，4，5，求ζ的分布就是求$P(\zeta=k)$，$k=0, 1, \cdots, 5$。为了说明问题具体，先求$P(\zeta=3)$。

因为3次命中可能出现在5次射击中的任何3次，所以5次射击中命中3次共有$\binom{5}{3}=10$种情况，因此，事件"$\zeta=3$"是下面10个互斥事件的和（A表示命中，$\overline{A}$表示未命中）

$$AAA\overline{A}\,\overline{A}\quad AA\overline{A}A\overline{A}\quad AA\overline{A}\,\overline{A}A\quad A\overline{A}AA\overline{A}\quad A\overline{A}A\overline{A}A$$
$$A\overline{A}\,\overline{A}AA\quad \overline{A}AAA\overline{A}\quad \overline{A}AA\overline{A}A\quad \overline{A}A\overline{A}AA\quad \overline{A}\,\overline{A}AAA$$

根据概率的加法性质，有

$$P(\zeta=3)=P(AAA\overline{A}\,\overline{A})+P(AA\overline{A}A\overline{A})+\cdots$$
$$+P(\overline{A}\,\overline{A}AAA)$$

由于5次射击是相互独立的，所以

$$P(AAA\overline{A}\,\overline{A})=P(A)\cdot P(A)\cdot P(A)\cdot P(\overline{A})\cdot P(\overline{A})$$
$$=(0.9)^3(0.1)^2$$

同理 $$P(AA\overline{A}A\overline{A})=P(AA\overline{A}\,\overline{A}A)=\cdots=P(\overline{A}\,\overline{A}AAA)$$
$$=(0.9)^3(0.1)^2$$

因此得 $P(\zeta=3)=10\cdot(0.9)^3(0.1)^2=\binom{5}{3}(0.9)^3(0.1)^2$.

用同样分析方法，可以得出，在5次射击中命中k次的概率是

$$P(\zeta=k)=\binom{5}{k}(0.9)^k(0.1)^{5-k},\ k=0, 1, 2, \cdots 5$$

一般地，在n重贝努利试验中A发生的次数ζ的概率分布$P(\zeta=k)$，类似于上例可如下求出。

由于事件A可能出现在n次试验中的任何次，所以，在n次试验中事件A出现k次，共有$\binom{n}{k}$种情况。因此，事件"$\xi=k$"等于$\binom{n}{k}$个"n次试验中，A在某指定的k次试验中发生，而在其余$n-k$次试验中不发生"这样事件的和。假定$P(A)=p$，则每一个这样事件的概率都等于p^kq^{n-k}，因而得到

$$P(\xi=k)=\binom{n}{k}p^kq^{n-k},\ k=0,1,2,\cdots,n \qquad (2.2.1)$$

通常把分布（2.2.1）叫做参数是$(n,\ p)$的二项分布，这样定义是因为$\binom{n}{k}p^kq^{n-k}$恰是二项式$(q+p)^n$的展开式

$$(q+p)^n=\binom{n}{0}p^0q^n+\binom{n}{1}pq^{n-1}+\binom{n}{2}p^2q^{n-2}+\cdots$$
$$+\binom{n}{k}p^kq^{n-k}+\cdots+\binom{n}{n}p^nq^0$$

中的第$k+1$项。由于$(q+p)^n=1$，所以

$$\sum_{k=0}^{n}P\ (\xi=k)=\sum_{k=0}^{n}\binom{n}{k}p^kq^{n-k}=(q+p)^n=1$$

这从另一角度说明了概率分布所具有的性质。

参数为$(n,\ p)$的二项分布用$b\ (n,\ p)$表示，其概率函数用$b\,(k,\ n,\ p)$表示，即$b(k,n,\ p)=\binom{n}{k}p^kq^{n-k}$。

例2.2.1　设每次射击的命中率是0.001，求“在5000次独立射击中命中两次以上”的概率。

解　用ξ表示在5000次射击中命中的次数，则所求概率可写成

$$P(\xi\geqslant2)=1-P(\xi<2)=1-P(\xi=0)-P(\xi=1)$$

显然ξ服从参数为（5000，0.001）的二项分布

$$P(\xi=k)=\binom{5000}{k}(0.001)^k(0.999)^{5000-k},\ k=0,1,2,\cdots,5000$$

所以　$P(\xi=0)=(0.999)^{5000}\doteq0.0067$

$P(\xi=1)=(5000)(0.001)(0.999)^{4999}\doteq0.0335$

因此，

$P(\xi\geqslant2)\doteq0.9598$

计算结果说明，尽管每次射击命中目标的可能性很小，但是多次射击，比如5000次，命中两次以上的概率就近似等于0.9598，至于命中一次以上的概率$P(\xi\geqslant1)$还要大，近似等于0.9933，很接近1，这实际上可以认为射击5000次几乎是能够命中目标的。

例2.2.2　已知某地区猪得某种病的概率是0.25，现有某种对

该病的新预防药。对12口猪进行用药试验，结果12口猪中都没有得该病。从这个结果我们对新预防药的效果能得出什么结论呢？

解　对12口猪进行试验，可看做是独立地进行12次试验。若药无效，则在每次试验中猪得病的概率都是 $p=0.25$，这时，12口猪中得病的头数 ξ 应当服从参数为（12，0.25）的二项分布，所以，12口猪都不得病”的概率是

$$P(\xi=0)=\binom{12}{0}(0.25)^0(0.75)^{12}=0.032$$

这说明，若药无效，则12口猪都不得病的可能性只有0.032，这个概率很小，在实际上不大可能发生。所以实际上可以认为药物是有效的。

例2.2.3　有甲、乙两盒同样的零件，每盒 N 个，当一工人装配机器时，他等可能地随机选一盒，从中取一个，他不断地装配机器，所以不断地取零件，求 $A=$ “当他第一次发现一盒已空时，另一盒还余 r 个零件”的概率。

解　先考虑事件“发现甲盒空，而乙盒还余 r 个”的概率。每取一个零件，只有两个可能：“取自甲盒”和“取自乙盒”，而且概率各为 $\frac{1}{2}$。事件“发现甲盒空，乙盒尚余 r 个”等价于事件“在 $2N-r$ 次取中，有 N 次取自甲而有 $N-r$ 次取自乙”与事件“在第 $2N-r+1$ 次取自甲”的积，根据二项分布的概率计算所求概率为

$$P=\binom{2N-r}{N-r}\left(\frac{1}{2}\right)^{N-r}\left(\frac{1}{2}\right)^{N}\left(\frac{1}{2}\right)=\binom{2N-r}{N-r}\left(\frac{1}{2}\right)^{2N-r+1}$$

事件“发现一盒空，而另一盒尚有 r 个”的概率就是上面所求概率的2倍

$$P(A)=\binom{2N-r}{N-r}\left(\frac{1}{2}\right)^{2N-r}$$

例2.2.4　设产品的次品率是0.005，求在任意的10000件产品中，（1）有40件次品的概率；（2）次品不多于70件的概率。

解　用 ξ 表示10000件产品中出现的次品数，则 ξ 服从参数为（10000，0.005）的二项分布，所以

（1）有40件次品的概率为

$$P(\zeta=40)=\binom{10000}{40}(0.005)^{40}(0.995)^{9960}$$

（2）次品不多于70件的概率为

$$P(\zeta\leqslant 70)=\sum_{k=0}^{70}P(\zeta=k)$$

$$=\sum_{k=0}^{70}\binom{10000}{k}(0.005)^{k}(0.995)^{10000-k}$$

计算例2.2.4的数值结果是很困难的，在下一节及第三章，我们将给出n很大时的二项分布的近似计算。

下面说明超几何分布与二项分布的关系。在例2.1.3中给出，若一批产品共N件，其中有M件废品，从中随机取出n（$n<N$）件，则n件中的废品数ξ服从超几何分布

$$P(\xi=k)=\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}\qquad \begin{array}{l}k\leqslant n\\ \\ k\leqslant M\end{array}\quad n-k\leqslant N-M$$

注意，这里在取n件产品时，每取一件都不再放回，因此每次取时，废品率都不一样，若在取n件产品时，每取完一个然后放回去，则每次取时，废品率都一样，为$\frac{M}{N}$，这时n件产品中的废品数ξ就服从二项分布

$$P(\xi=k)=\binom{n}{k}\left(\frac{M}{N}\right)^{k}\left(\frac{N-M}{N}\right)^{n-k}$$

按两种不同的抽取法，所得概率不一样。但当N很大时，既使是不放回地抽取产品，每次抽得废品的概率也变化不大，因此，当N很大时，超几何分布应与二项分布很接近，这一事实可用下面定理说明。

定理2.2.1　若当$N\longrightarrow\infty$时，$\frac{M}{N}\longrightarrow p>0$，$\frac{N-M}{N}\longrightarrow q=1-p>0$，则有

$$\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}\longrightarrow\binom{n}{k}p^{k}q^{n-k}$$

证明

$$\frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}}=\frac{M!}{k!(M-k)!}$$

$$\cdot\frac{(N-M)!}{[(N-M)-(n-k)]!(n-k)!}\cdot\frac{n!(N-n)!}{N!}$$

$$=\frac{n!}{k!(n-k)!}\cdot\frac{M!}{(M-k)!}\cdot\frac{(N-M)!}{[(N-M)-(n-k)]!}\cdot\frac{(N-n)!}{N!}$$

$$=\binom{n}{k}\frac{M(M-1)\cdots(M-k+1)}{N(N-1)\cdots}$$

$$\cdot\frac{(N-M)(N-M-1)\cdots(N-M-n+k+1)}{\cdots(N-n+1)}$$

$$=\binom{n}{k}\frac{\frac{M}{N}\left(\frac{M}{N}-\frac{1}{N}\right)\cdots\left(\frac{M}{N}-\frac{k-1}{N}\right)\cdot}{\left(1-\frac{1}{N}\right)\cdots}$$

$$\cdot\frac{\left(1-\frac{M}{N}\right)\left(1-\frac{M}{N}-\frac{1}{N}\right)\cdots\left(1-\frac{M}{N}-\frac{n-k-1}{N}\right)}{\cdots\left(1-\frac{n-1}{N}\right)}$$

$$\xrightarrow[N\to\infty]{}\binom{n}{k}p^k(1-p)^{n-k}$$

现在我们分析一下，二项分布随机变量的概率函数 $b(k,n,p)$，对于固定的 n，p，当值 k 变化时它是怎样变化的。

对 k 的相近两个值比较一下看。

$$\frac{b(k,n,p)}{b(k-1,n,p)}=\frac{\binom{n}{k}p^kq^{n-k}}{\binom{n}{k-1}p^{k-1}q^{n-k+1}}=\frac{(n-k+1)p}{kq}$$

$$=\frac{(n+1)p-kp}{kq}=\frac{(n+1)p-k(1-q)}{kq}$$

$$=\frac{kq+(n+1)p-k}{kq}=1+\frac{(n+1)p-k}{kq}$$

（a） 若$k<(n+1)p$，则 $b(k, n, p)>b(k-1, n, p)$，说明当 $k<(n+1)p$ 时，函数单调上升。

（b） 若 $k>(n+1)p$，则$b(k, n, p)<b(k-1, n, p)$，说明当$k>(n+1)p$时，函数单调下降。

（c） 若 $k=(n+1)p$，则$b(k, n, p)=b(k-1, n, p)$，如果$(n+1)p$为正整数，则函数在 $k=(n+1)p$及$(n+1)p-1$时达到最大值；但$(n+1)p$ 不一定是正整数，这时必存在满足

$$(n+1)p-1<k<(n+1)p$$

的正整数k，由于$k+1>(n+1)p$ 及$k<(n+1)p$，根据(a)，(b)，则有

$$b(k+1, n, p)<b(k, n, p)$$

$$b(k, n, p)>b(k-1, n, p)$$

所以，这时$k=[(n+1)p]$时函数达到最大值。使$b(k, n, p)$取最大值的 k 称为最可能成功次数，$b(k, n, p)$的最大值称为$b(k, n, p)$的中心项。总之，函数$b(k, n, p)$当 k 由 0 变到 n 时，开始单调上升，当$k=[(n+1)p]$($(n+1)p$为正整数时，k 取 $(n+1)p$及$(n+1)p-1$)时达到最大值，然后又单调下降。

例 一批产品废品率为0.01，从中随机取出200件，问200件中废品数是多少的概率最大？

解 这里$n=200$，$p=0.01$，$(n+1)p=2.01$

所以废品数为 2 的可能性最大。概率为

$$\binom{200}{2}(0.01)^2(0.99)^{198}\doteq 0.272$$

§2.3 普阿松分布

（一）普阿松定理

上节求出了二项分布的计算公式（2.2.1），由于在实际中遇到的二项分布常常是n很大，这样，（2.2.1）用起来就很困难，比如，在上节的例2.2.1中，如果要求计算

$$P(\xi=10)=\binom{5000}{10}(0.001)^{10}(0.999)^{4990}$$

这很麻烦。因此有必要解决当n很大时二项分布的近似计算问题。下面先对n很大，而P又很小的情况加以讨论。

普阿松定理　对于二项分布$b(k,n,p_n)$，若当$n\longrightarrow\infty$时，$\lambda_n=np_n\longrightarrow\lambda$（常数），则

$$b(k,n,p_n)=\binom{n}{k}p_n^k(1-p_n)^{n-k}\longrightarrow\frac{\lambda^k}{k!}e^{-\lambda}$$

$$k=0,\ 1,\ 2,\ \cdots,n \tag{2.3.1}$$

（这里把p写为p_n是表示p与试验次数n有关。$np_n\longrightarrow\lambda$表示$n\longrightarrow\infty$时，$p_n\longrightarrow 0$）

证明

$$\binom{n}{k}p_n^k(1-p_n)^{n-k}$$

$$=\frac{n(n-1)(n-2)\cdots(n-k+1)}{k!}\left(\frac{\lambda_n}{n}\right)^k\left(1-\frac{\lambda_n}{n}\right)^{n-k}$$

$$=\frac{\lambda_n^k}{k!}\cdot\frac{n}{n}\cdot\frac{n-1}{n}\cdots\frac{n-k+1}{n}\left(1-\frac{\lambda_n}{n}\right)^{n-k}$$

对于任一固定的k，有

$$\lim_{n\to\infty}\frac{n}{n}\cdot\frac{n-1}{n}\cdot\frac{n-2}{n}\cdots\frac{n-k+1}{n}=1$$

而

$$\left(1-\frac{\lambda_n}{n}\right)^{n-k}=\left(1-\frac{\lambda_n}{n}\right)^n\left(1-\frac{\lambda_n}{n}\right)^{-k}$$

$$=\left[\left(1-\frac{\lambda_n}{n}\right)^{-\frac{n}{\lambda_n}}\right]^{-\lambda_n}\left(1-\frac{\lambda_n}{n}\right)^{-k}$$

由于　$\frac{\lambda_n}{n}=p_n\xrightarrow[n\to\infty]{}0$，所以

$$\lim_{n\to\infty}\left[\left(1-\frac{\lambda_n}{n}\right)^{-\frac{n}{\lambda_n}}\right]^{-\lambda_n}\left(1-\frac{\lambda_n}{n}\right)^{-k}=e^{-\lambda}$$

所以当 $n\to\infty$ 时，有

$$b(k,\ n,\ p_n)\longrightarrow\frac{\lambda^k}{k!}e^{-\lambda}$$

注意，在导出（2.3.1）时用到了条件：$p_n\xrightarrow[n\to\infty]{}0$，因此在实际应用（2.3.1）时，只有在 n 很大，p 很小，而 $\lambda=np$ 大小适中时才比较适用。通常 $np\leqslant5$ 比较合适。这时可用下面的近似公式

$$\binom{n}{k}p^kq^{n-k}\doteq\frac{(np)^k}{k!}e^{-(np)}$$

（2.3.1）右端的函数 $\frac{\lambda^k}{k!}e^{-\lambda}$ 有专门的表可查，对于已知的 λ 和 k 值，表中给出了函数的值。

在定理的条件满足时，用（2.3.1）作二项分布的近似计算，其近似程度完全符合实际需要，这一点可以从下面事实看出。下表对参数为（10，0.1），（20，0.05），（40，0.025）以及（100，0.01）的二项分布比较了普阿松逼近的程度。

k	按二项分布的公式计算				普阿松近似
	$n=10$ $p=0.1$	$n=20$ $p=0.05$	$n=40$ $p=0.025$	$n=100$ $p=0.01$	$\lambda=1$
0	0.349	0.358	0.363	0.366	0.368
1	0.385	0.377	0.372	0.370	0.368
2	0.194	0.189	0.186	0.185	0.184
3	0.057	0.060	0.060	0.061	0.061
4	0.011	0.013	0.014	0.015	0.015
>4	0.004	0.003	0.005	0.003	0.004

从上表看到当 n 不太大时，逼近程度差一些，当 n 较大时，逼近程度是很好的。

例2.3.1 用普阿松定理，求上节例2.2.1中 $P(\zeta\geqslant2)$，并计算 $P(\zeta=10)$

解 在这个问题中 $n=5000$，$p=0.001$，$\lambda=5$。查函数 $\frac{\lambda^k}{k!}e^{-\lambda}$

的数值表，当 $k=0$时，函数值为 0.00673，当 $k=1$ 时，函数值为0.03369。于是得

$$P(\zeta\geqslant 2)=1-P(\zeta=0)-P(\zeta=1)\doteq 1-0.00673-0.03369$$
$$=0.95958$$

此结果与用二项分布计算公式算得结果极为近似。

再对 $k=10$查表得

$$P(\zeta=10)\doteq 0.018$$

例2.3.2 设螺丝钉的废品率是0.01，螺丝钉每100多个装一盒。问在盒中应装多少个螺丝钉才能使其中"合格品数≥100"的概率≥0.95?

解 设$n=100+x$ 是应装的螺丝钉数，用 ζ 表示 $n=100+x$个螺丝钉中的废品数，则"合格品数≥100"等价于"$\zeta\leqslant x$"，所以所求的x 应满足下面不等式

$$P(\zeta\leqslant x)=P(\zeta=0)+P(\zeta=1)+\cdots+P(\zeta=x)\geqslant 0.95$$

根据普阿松定理，上面不等式可近似写成

$$\sum_{k=0}^{x}\frac{\lambda^k}{k!}e^{-\lambda}\geqslant 0.95$$

由于 x 必是很小的正整数，所以$\lambda=np=(100+x)(0.01)\doteq 1$。对于$\lambda=1$，查普阿松分布表可知

当 $x=1$ 时 不等式左边等于

$$0.36788+0.36788=0.73576$$

当 $x=2$ 时 不等式左边等于

$$0.73576+0.18394\doteq 0.9197$$

当 $x=3$ 时 不等式左边等于

$$0.9197+0.06131=0.98101$$

已超过0.95。所以得出结论：一盒中装103个能使其中"合格品数≥100"的概率≥0.95。

下表说明，当 p 较大时，用普阿松定理作近似计算的误差较大。对这种情况如何进行近似计算将在下一章加以讨论。

k	按二项分布计算 $n=40 \quad p=0.25$	普阿松近似 $\lambda=np=10$	k	按二项分布计算 $n=40 \quad p=0.25$	普阿松近似 $\lambda=np=10$
0	0.0000	0.0000	11	0.1312	0.1137
1	0.0001	0.0005	12	0.1057	0.0948
2	0.0009	0.0023	13	0.0759	0.0729
3	0.0037	0.0076	14	0.0488	0.0521
4	0.0113	0.0189	15	0.0282	0.0347
5	0.0272	0.0378	16	0.0147	0.0217
6	0.0529	0.0631	17	0.0070	0.0128
7	0.0857	0.0901	18	0.0029	0.0071
8	0.1179	0.1126	19	0.0011	0.0037
9	0.1398	0.1251	20	0.0004	0.0019
10	0.1444	0.1251	大于20	0.0002	0.0015

（二）普阿松分布

由普阿松定理已知，对二项分布，当$n \longrightarrow \infty$时，若 $np_n \to \lambda$，则

$$b(k, n, p_n) \longrightarrow \frac{\lambda^k}{k!}e^{-\lambda}, \quad k=0, 1, 2, \cdots, n$$

这里$n \longrightarrow \infty$，就意味着随机变量的可能取值的个数趋于无穷。因此，自然容易想到，函数$\frac{\lambda^k}{k!}e^{-\lambda}$ 本身可不可能是一个取值为 0，1，2，…的随机变量的概率分布呢？这一点由下面事实得到肯定回答。

对任一$\lambda > 0$

$$\frac{\lambda^k}{k!}e^{-\lambda} > 0$$

而且 $\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}e^{-\lambda}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^k}{k!}=e^{-\lambda}e^{\lambda}=1$。

这样，我们得到了一个可能值为所有非负整数 $k=0, 1, 2, \cdots$ 的随机变量的概率分布

$$P(\xi=k)=\frac{\lambda^k}{k!}e^{-\lambda},\quad k=0,\ 1,\ 2,\ \cdots \tag{2.3.2}$$

(2.3.2) 叫做参数为λ的普阿松分布，有时用$p(\lambda)$表示。它是概率论中最重要的概率分布之一，上面我们是把它作为二项分布的极限导出的，因此，可近似地把它看作是一个概率很小的事件，在大量试验中出现的次数的概率分布。但是，它的重要意义还不仅仅是能作二项分布的近似计算，而在实际中有很多随机变量本身就是服从普阿松分布的。比如，在一段时间里，电话交换台来到的呼唤数；公共汽车站来到的乘客数；一块地里发生的害虫数；一棵苹果树上被虫食的苹果数；放射性物质落到某区域里的质点数；某操作系统发生故障的次数等等都是服从普阿松分布的。

为什么说这些随机变量是服从普阿松分布的？一个随机变量服从普阿松分布的条件是什么呢？

粗略地说，如果某事件的发生与否（如呼唤的来到，乘客的来到，害虫的来到，…，故障的发生，等等）在时间上是随机的，那么，在一定时间内该事件发生的次数ξ，在一定条件下就服从普阿松分布。

下面我们结合电话交换台来到的呼唤数简略地说明一下这个问题。

现在我们来求电话交换台在时间区间$(0,t)$内来到的呼唤数ξ的概率分布$P(\xi=k)$，$k=0, 1, 2, \cdots$。我们把时间区间设想分割为n等分，使每一等分长Δt都很小，并假定

（1）对于每个小区间来说，来到两次以上呼唤的概率很小，可近似地看作为0，即在每个小区间里只能出现来到一次呼唤或没有呼唤两种情况。

（2）各小区间里是否来到呼唤是相互独立的。

在这两条假定下，在时间区间$(0, t)$里来到k次呼唤这一事件，可近似地看作是在$(0,t)$中的n个独立小区间中恰有k个小区间各来到一次呼唤，所以它的概率近似地等于

$$\binom{n}{k}p_n^k q_n^{n-k}, \quad k=0, 1, 2, \cdots, n$$

其中p_n表示在每个小区间里来到一次呼唤的概率。如果我们把$(0, t)$无限细分，即n充分大，自然应当得到$P(\xi=k)$的精确式，因此，根据普阿松定理，我们有

$$P(\xi=k)=\frac{\lambda^k}{k!}e^{-\lambda}, \quad k=0, 1, 2, \cdots$$

其中 $\lambda=\lim\limits_{n\to\infty}np_n$

§2.4 随机变量函数的分布

设ξ为一离散型随机变量，$f(x)$是定义在随机变量ξ的一切可能值x的集合上的函数，若当随机变量ξ取值x时，随机变量η就取值$y=f(x)$，则称η为随机变量ξ的函数，记作$\eta=f(\xi)$，显然η仍是一个离散型随机变量。同样可定义两个随机变量ξ_1，ξ_2的函数$\eta=f(\xi_1, \xi_2)$。这一节要解决的问题是，如何根据已知随机变量ξ的分布，求$\eta=f(\xi)$的分布，或者已知ξ_1，ξ_2的联合分布，求$\eta=f(\xi_1, \xi_2)$的分布。

下面先讨论单个随机变量函数的分布，先看一个例子。

例2.4.1 设已知ξ的分布为

$$\begin{pmatrix} -1, & 0, & +1 \\ 0.2, & 0.7, & 0.1 \end{pmatrix}$$

求 $\eta=\xi^2$的分布。

先确定η的可能取值，可知为0，1。由于

$$(\eta=0)=(\xi=0)$$

$$(\eta=1)=(\xi=-1)\cup(\xi=+1)$$

所以

$$P(\eta=0)=P(\xi=0)=0.7$$

$$P(\eta=1)=P(\xi=-1)+P(\xi=+1)=0.2+0.1=0.3$$

因此，η的分布列为

$$\begin{pmatrix} 0, & 1 \\ 0.7, & 0.3 \end{pmatrix}$$

一般地，若ξ是一个离散型随机变量，可能取值为x_1，x_2，…，$\eta=f(\xi)$的可能取值为y_1，y_2，…，则

$$(\eta=y_k)=\sum_{f(x_i)=y_k}(\xi=x_i)$$

等式右端是对所有使 $f(x_i)=y_k$的x_i求和，所以

$$P(\eta=y_k)=\sum_{f(x_i)=y_k}P(\xi=x_i)$$

例2.4.2　已知ξ服从二项分布

$$P(\xi=k)=\binom{n}{k}p^k q^{n-k},\quad k=0,1,2,\cdots,n$$

试求（1）$\eta=a\xi+b$及（2）$\eta=\xi^2$的分布，a，b为任意二实数。

解　（1）由于η的可能取值为$ak+b$，$k=0$，1，2，…，n，而$(\eta=ak+b)=(\xi=k)$

所以

$$P(\eta=ak+b)=P(\xi=k)=\binom{n}{k}p^k q^{n-k},\ k=0,1,2,\cdots,n$$

（2）η的可能取值为k^2，而

$$(\eta=k^2)=(\xi=k)$$

所以

$$P(\eta=k^2)=P(\xi=k)=\binom{n}{k}p^k q^{n-k},\ k=0,1,2,\cdots,n$$

其次考虑两个随机变量的函数的分布。

设(ξ,η)为一个二维离散型随机变量，ξ，η的可能取值分别为x_1，x_2，…，y_1，y_2，…，$\zeta=f(\xi,\eta)$的可能取值为z_k，则

$$(\zeta=z_k)=[f(\xi,\eta)=z_k]=\sum_{f(x_i,y_j)=z_k}(\xi=x_i\ \eta=y_j)$$

等式右端是对所有使$f(x_i,y_j)=z_k$的(x_i,y_j)求和，所以

$$P(\zeta = z_k) = \sum_{f(x_i, y_j) = z_k} P(\xi = x_i, \eta = y_j)$$

例2.4.3　设ξ_1，ξ_2相互独立且服从相同的两点分布$\begin{pmatrix} 0, & 1 \\ q, & p \end{pmatrix}$，求$\xi_1+\xi_2$的分布。

解　我们知道，n次独立重复贝努利试验中，A出现的次数ζ等于在每次试验中A出现次数的和，即

$$\zeta = \mu_1 + \mu_2 + \cdots + \mu_n$$

μ_i为第i次试验中A出现的次数，它服从两点分布$\begin{pmatrix} 0, & 1 \\ q, & p \end{pmatrix}$，而$\zeta$服从二项分布。所以可知$(\xi_1+\xi_2)$服从参数为$(2, p)$的二项分布。

现在我们再用上述方法直接求一下。

$\xi_1+\xi_2$的可能取值显然为0，1，2，由于

$$(\xi_1+\xi_2=0) = (\xi_1=0)\cdot(\xi_2=0)$$

$$(\xi_1+\xi_2=1) = (\xi_1=0)\cdot(\xi_2=1) \cup (\xi_1=1)\cdot(\xi_2=0)$$

$$(\xi_1+\xi_2=2) = (\xi_1=1)\cdot(\xi_2=1)$$

所以

$$P(\xi_1+\xi_2=0) = q^2$$

$$P(\xi_1+\xi_2=1) = 2pq$$

$$P(\xi_1+\xi_2=2) = p^2$$

例2.4.4　设随机变量ξ_1，ξ_2相互独立，且分别服从参数为λ_1，λ_2的普阿松分布，试求（1）$\xi_1+\xi_2$的分布；（2）条件分布$P(\xi_1=k | \xi_1+\xi_2=n)$。

解　（1）$\xi_1+\xi_2$的可能取值为0，1，2，…，

$$\begin{aligned} P(\xi_1+\xi_2=k) &= P(\xi_1=0,\xi_2=k)+P(\xi_1=1,\xi_2=k-1)+ \\ &\quad \cdots + P(\xi_1=k, \xi_2=0) \\ &= \sum_{i=0}^{k} P(\xi_1=i \quad \xi_2=k-i) \\ &= \sum_{i=0}^{k} P(\xi_1=i)\cdot P(\xi_2=k-i) \end{aligned}$$

$$= \sum_{i=0}^{k} \frac{\lambda_1^i}{i!} e^{-\lambda_1} \cdot \frac{\lambda_2^{k-i}}{(k-i)!} e^{-\lambda_2}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{k!} \sum_{i=0}^{k} \frac{k!}{(k-i)!\,i!} \lambda_1^i \lambda_2^{k-i}$$

$$= \frac{(\lambda_1+\lambda_2)^k}{k!} e^{-(\lambda_1+\lambda_2)}$$

可知 $\xi_1+\xi_2$ 服从参数为 $\lambda_1+\lambda_2$ 的普阿松分布。

(2) $$P(\xi_1=k \mid \xi_1+\xi_2=n) = \frac{P(\xi_1=k, \xi_1+\xi_2=n)}{P(\xi_1+\xi_2=n)}$$

$$= \frac{P(\xi_1=k, \xi_2=n-k)}{P(\xi_1+\xi_2=n)}$$

$$= \frac{P(\xi_1=k) \cdot P(\xi_2=n-k)}{P(\xi_1+\xi_2=n)}$$

$$= \frac{\dfrac{\lambda_1^k}{k!} e^{-\lambda_1} \cdot \dfrac{\lambda_2^{n-k}}{(n-k)!} e^{-\lambda_2}}{\dfrac{(\lambda_1+\lambda_2)^n}{n!} e^{-(\lambda_1+\lambda_2)}}$$

$$= \binom{n}{k} \left(\frac{\lambda_1}{\lambda_1+\lambda_2} \right)^k \left(\frac{\lambda_2}{\lambda_1+\lambda_2} \right)^{n-k}$$

$$k = 0, 1, 2, \cdots$$

可知它服从二项分布。

§2.5 离散型随机变量的数字特征

在实际中经常需要了解一个随机变量的某一方面的特征，比如了解一个随机变量平均取什么值，或者了解一个随机变量取值的分散程度，等等。能用数字表示的特征叫数字特征。由于数字特征的表现形式简单，所以在实际中了解一个随机变量的数字特征比起了解随机变量的概率分布容易得多。而且对实际中许多问题只要知道

随机变量的某些数字特征也就足够了．因此研究随机变量的数字特征很重要．从理论与应用两方面考虑，最重要的数字特征有两个：均值与方差．

（一）均　　值

粗略地说，均值就是随机变量的平均取值．比如，电话交换台在一段时间内接到的呼唤数是一个随机变量，它的均值就是平均起来交换台在一段时间内接到的呼唤数．在给出均值的数学定义之前，先看下面的例子．

一个车间共有5台机床，这些机床由于各种原因时而工作时而停顿，因此任一时刻工作着的机床数是一个随机变量．为了精确估计车间的电力负荷，我们需要知道车间中同时工作着的机床的平均数．

假定我们作了20次观察，结果如表2.5.1．

结果说明，在20次观察中，有1次“有2台工作，”有2次“有3台工作”，有6次“有4台工作”，有11次“有5台工作”，“机床都不工作”及“有1台工作”的情况没出现，在20次观察中，工作着的机床总数为

$$0\times0+1\times0+2\times1+3\times2+4\times6+5\times11=87$$

表2.5.1

工作着的机床数	0	1	2	3	4	5
频　数	0	0	1	2	6	11
频　率	0	0	$\frac{1}{20}$	$\frac{2}{20}$	$\frac{6}{20}$	$\frac{11}{20}$

因此，车间中同时工作着的机床的平均数为

$$\frac{0\times0+1\times0+2\times1+3\times2+4\times6+5\times11}{20}$$

$$= 0 \times \frac{0}{20} + 1 \times \frac{0}{20} + 2 \times \frac{1}{20} + 3 \times \frac{2}{20} + 4 \times \frac{6}{20} + 5 \times \frac{11}{20}$$

$$= 4.35$$

所求的平均数实际上就是随机变量所能取值与对应的频率乘积之和．由于这个平均数是由观察得来的，因此它具有偶然性，这种偶然性主要反映在频率上．如果我们用概率来代替频率，就消除了这种偶然性而从本质上反映了随机变量的平均值．因此，对离散型随机变量定义均值如下．

定义2.5.1　设随机变量ξ的分布列为

$$\begin{pmatrix} x_1, & x_2, & \cdots, & x_k, & \cdots \\ p_1, & p_2, & \cdots, & p_k, & \cdots \end{pmatrix}$$

若级数 $$\sum_{k=1}^{\infty} x_k p_k \tag{2.5.1}$$

绝对收敛，则定义级数（2.5.1）为ξ的均值（或叫数学期望）．用$E\xi$表示．

一个随机变量的均值，按其实际意义，应当是一个确定的值，因此它不应因改变级数项的顺序而改变．这就是在均值定义中要求（2.5.1）绝对收敛的道理．

例2.5.1　常数c的均值等于它本身

$$Ec = c \cdot 1 = c$$

例2.5.2　两点分布随机变量的均值为

$$E\xi = 0 \cdot q + 1 \cdot p = p$$

例2.5.3　服从二项分布

$$P(\xi = k) = \binom{n}{k} p^k q^{n-k}, \quad k = 0, 1, \cdots, n$$

的随机变量的均值为

$$E\xi = \sum_{k=0}^{n} k \cdot \binom{n}{k} p^k (1-p)^{n-k}$$

$$= \sum_{k=1}^{n} k \frac{n!}{k!(n-k)!} p^k q^{n-k}$$

$$=\sum_{k=1}^{n} np\frac{(n-1)!}{(k-1)!([n-1)-(k-1)]!}p^{k-1}q^{(n-1)-(k-1)}$$

$$=np\sum_{k=1}^{n}\binom{n-1}{k-1}p^{k-1}q^{(n-1)-(k-1)}$$

$$=np\sum_{k=0}^{n-1}\binom{n-1}{k}p^{k}q^{(n-1)-k}$$

$$=np(p+q)^{n-1}=np$$

如果已知每次射击的命中率为$p=0.9$，那么在50次射击中命中数ξ的均值就是$50\times0.9=45$，这个结果直观上是容易理解的。

例2.5.4　服从普阿松分布

$$P(\xi=k)=\frac{\lambda^k}{k!}e^{-\lambda},\qquad k=0,1,2,\cdots$$

的随机变量的均值为

$$E\xi=\sum_{i=1}^{\infty}x_ip_i=\sum_{k=0}^{\infty}k\cdot\frac{\lambda^k}{k!}e^{-\lambda}=\sum_{k=1}^{\infty}\frac{\lambda^k}{(k-1)!}e^{-\lambda}$$

$$=\lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}=\lambda e^{-\lambda}e^{\lambda}=\lambda.$$

例2.5.5　服从几何分布

$$P(\xi=k)=q^{k-1}p,\qquad k=1,2,\cdots$$

的随机变量的均值为

$$E\xi=\sum_{k=1}^{\infty}k\cdot q^{k-1}p=p(1+2q+3q^2+\cdots)$$

$$=p(q+q^2+q^3+\cdots)'$$

$$=p\left(\frac{q}{1-q}\right)'=p\frac{1}{(1-q)^2}=\frac{1}{p}$$

关于随机变量函数的均值，我们有如下定理：

定理2.5.1　（1）若ξ为一离散型随机变量，其分布为$P(\xi=x_i)=p(x_i)$，$f(x)$为一连续函数，若级数$\sum_{i=1}^{\infty}f(x_i)p(x_i)$绝对收敛，则随机变量$f(\xi)$的均值为

$$Ef(\xi)=\sum_{i=1}^{\infty}f(x_i)p(x_i)$$

（2）若 (ξ,η) 为二维随机变量，联合分布为 $P(\xi=x_j,\ \eta=y_k)=p(x_j,y_k)$，$f(x,\ y)$为二元连续函数，若级数$\sum\limits_{j}\sum\limits_{k}f(x_j,\ y_k)p(x_j,\ y_k)$绝对收敛，则随机变量$\zeta=f(\xi,\ \eta)$的均值为

$$E\zeta=\sum_{j}\sum_{k}f(x_j,\ y_k)p(x_j,\ y_k)$$

（证明从略）

均值有下面几个性质

性质 1　对任一常数 c，有

$$E(c\xi)=cE\xi$$

这性质由定义直接可得。

性质 2　对任意的 n 个随机变量 ξ_1，ξ_2，…，ξ_n，若相应的均值存在，则

$$E(\xi_1+\xi_2+\cdots+\xi_n)=E\xi_1+E\xi_2+\cdots+E\xi_n, \tag{2.5.2}$$

就是说，随机变量和的均值等于各随机变量均值的和。

证明　只对两个随机变量 ξ，η 的情况来证明就够了。

设 ξ，η，$(\xi,\ \eta)$ 的概率分布分别为

$$P\{\xi=x_j\}=g(x_j),$$

$$P\{\eta=y_k\}=h(y_k),$$

$$P(\xi=x_j,\ \eta=y_k)=p(x_j,\ y_k)$$

$$j,\ k=1,\ 2,\ \cdots$$

则

$$\begin{aligned}E\xi+E\eta&=\sum_{j}x_jg(x_j)+\sum_{k}y_kh(y_k)\\&=\sum_{j}x_j\sum_{k}p(x_j,y_k)+\sum_{k}y_k\sum_{j}p(x_j,y_k)\\&=\sum_{j,k}x_jp(x_j,y_k)+\sum_{j,k}y_kp(x_j,\ y_k)\\&=\sum_{j,k}(x_j+y_k)p(x_j,y_k)\end{aligned}$$

后式恰为随机变量$\xi+\eta$的均值$E(\xi+\eta)$.

性质3　若相互独立的随机变量ξ与η的均值存在，则

$$E(\xi\eta)=E\xi\cdot E\eta \tag{2.5.3}$$

就是说，在独立的条件下，随机变量积的均值等于各随机变量均值的积

证明　$$E(\xi\eta)=\sum_{j,k}x_jy_kp(x_j,y_k)$$

$$=\sum_{j,k}x_jy_kg(x_j)h(y_k)$$

$$=\sum_{j,k}\{x_jg(x_j)\}\cdot\{y_k\cdot h(y_k)\}$$

$$=\left\{\sum_j x_jg(x_j)\right\}\cdot\left\{\sum_k y_kh(y_k)\right\}=E\xi\cdot E\eta$$

利用均值的性质，可用另一方法求二项分布的均值.

若在每次试验中事件A出现的概率为p，以ζ表示在n次独立试验中事件A出现的次数，而以μ_1，μ_2，…，μ_n分别表示在第一次，第二次，…，第n次试验中A出现的次数.显然，在n次试验中A出现的次数是A在各次试验中出现次数的和，即

$$\zeta=\mu_1+\mu_2+\cdots+\mu_n$$

因为在各次试验中A出现的次数，只有两个可能值：0与1，而取1的概率为p。所以μ_i的分布列为

$$\begin{pmatrix}0, & 1\\ q, & p\end{pmatrix}$$

因而 $E\mu_i=p, i=1, 2, \cdots, n$，根据性质2则得

$$E\zeta=E\mu_1+E\mu_2+\cdots+E\mu_n=np$$

与用均值定义直接求的结果一致.

（二）方　　差

在不少的实际问题里，不仅需要知道随机变量的均值，而且还

需要知道随机变量取值与均值的偏差程度．比如，考虑两名射手打靶射击水平，不仅要看他们各自平均击中环数，而且还应当看他们击中环数的摆动程度．假定两名射手甲与乙各射击了5次，所得环数如下：

甲	乙
4，8，7，10，6，	7，7，8，7，6

平均环数都是7环，但击中环数与平均环数的偏离程度确不一样．从摆动的情况看乙比甲好．随机变量的这一特性用均值反映不出来．因此应当引进一个数，用以刻划随机变量对它的均值的偏离程度．

对于上述例子可以这样做，先求每个实际取值与平均值的差的平方：$(4-7)^2$，$(8-7)^2$，…，$(6-7)^2$ 及 $(7-7)^2$，$(7-7)^2$，…，$(6-7)^2$，然后分别求它们的平均，

$$甲\quad \frac{1}{5}(9+1+0+9+1)=4$$

$$乙\quad \frac{1}{5}(0+0+1+0+1)=2/5=0.4$$

由于$0.4<4$，所以乙比甲稳．

在这里为什么用实际取值与均值的差的平方，而不用差本身，这是因为差本身可能有正值也可能有负值，正值与负值相互抵消，因而反映不出偏离程度，用差的平方就避免了这种情况．

由于这个例子启发，想到可用$(\xi-E\xi)^2$的均值$E(\xi-E\xi)^2$描述ξ对其均值$E\xi$的偏离程度，因而给出方差定义如下．

定义2.5.2　对于随机变量ξ，若$(\xi-E\xi)^2$的均值$E(\xi-E\xi)^2$存在，则称其为ξ的方差，用$D\xi$表示，即

$$D\xi=E(\xi-E\xi)^2 \qquad (2.5.5)$$

根据(2.5.5)可知

$$D\xi=E(\xi-E\xi)^2=\sum_{k=1}^{\infty}(x_k-E\xi)^2P(\xi=x_k)$$

方差$D\xi$的开方$\sqrt{D\xi}$叫做ξ的标准差(或均方差)．

下面推导一个常用的计算方差的公式

$$D\xi = E(\xi - E\xi)^2 = E[\xi^2 - 2\xi \cdot E\xi + (E\xi)^2]$$
$$= E\xi^2 - 2E\xi \cdot E\xi + (E\xi)^2 = E\xi^2 - (E\xi)^2 \qquad (2.5.6)$$

在计算方差时，有时用(2.5.6)比用(2.5.5)方便。

方差有下面几个性质

性质1　对任一常数 c，有 $D(c\xi) = c^2 D\xi$。这性质由定义直接可得。

性质2　对相互独立的随机变量 $\xi_1, \xi_2, \cdots, \xi_n$，若相应的方差存在，则

$$D(\xi_1 + \xi_2 + \cdots + \xi_n) = D\xi_1 + D\xi_2 + \cdots + D\xi_n \qquad (2.5.7)$$

就是说，相互独立的随机变量和的方差等于各随机变量方差的和。

证明　只需对两个随机变量 ξ，η 的情形进行证明就够了。

$$D(\xi + \eta) = E(\xi + \eta)^2 - [E(\xi + \eta)]^2$$
$$= E(\xi^2 + 2\xi\eta + \eta^2) - [(E\xi)^2 + (E\eta)^2 + 2(E\xi)(E\eta)]$$
$$= [E\xi^2 - (E\xi)^2] + [E\eta^2 - (E\eta)^2]$$
$$+ 2[E\xi\eta - (E\xi)(E\eta)]$$

由于 ξ 与 η 独立，所以，$E\xi\eta = (E\xi)(E\eta)$

因而有

$$D(\xi + \eta) = D\xi + D\eta$$

性质3　若 $c \neq E\xi$,，则

$$D\xi < E(\xi - c)^2$$

证明

$$D\xi = E(\xi - E\xi)^2 = E[(\xi - c) - (E\xi - c)]^2$$
$$= E(\xi - c)^2 - 2(E\xi - c)^2 + (E\xi - c)^2$$
$$= E(\xi - c)^2 - (E\xi - c)^2 < E(\xi - c)^2 \qquad (2.5.8)$$

例2.5.6　常数 c 的方差等于0

$$Dc = (c - c)^2 \cdot 1 = 0$$

这结果是直观的，因为常数对它自己本来就无任何偏差可言。

例2.5.7　两点分布随机变量的方差是pq

$$D\xi=(0-p)^2q+(1-p)^2p=p^2q+q^2p=pq(p+q)=pq$$

也可用（2.5.6）来计算

$$E\xi^2=0^2q+1^2\cdot p=p$$

$$D\xi=E\xi^2-(E\xi)^2=p-p^2=p(1-p)=pq$$

例2.5.8　求服从二项分布

$$P(\xi=k)=\binom{n}{k}p^kq^{n-k},\qquad k=0,\ 1,\ 2,\ \cdots,\ n$$

的随机变量的方差。

用$\mu_k(k=1,2,\cdots,n)$表示在第k次试验中事件A出现的次数，则在n次试验中事件A出现的次数

$$\xi=\mu_1+\mu_2+\cdots+\mu_n$$

由于μ_k相互独立，根据方差的性质2及例2.5.7，则得

$$D\xi=D\mu_1+D\mu_2+\cdots+D\mu_n=npq$$

例2.5.9　求服从普阿松分布

$$P(\xi=k)=\frac{\lambda^k}{k!}e^{-\lambda},\qquad k=0,1,2,\cdots$$

的随机变量的方差。

根据方差计算公式（2.5.6）先求$E\xi^2$

$$E\xi^2=\sum_{k=0}^{\infty}k^2\frac{\lambda^k}{k!}e^{-\lambda}=\sum_{k=1}^{\infty}k\frac{\lambda^k}{(k-1)!}e^{-\lambda}$$

$$=\sum_{k=1}^{\infty}[(k-1)+1]\frac{\lambda^k}{(k-1)!}e^{-\lambda}$$

$$=\sum_{k=1}^{\infty}(k-1)\cdot\frac{\lambda^k}{(k-1)!}e^{-\lambda}+\sum_{k=1}^{\infty}\frac{\lambda^k}{(k-1)!}e^{-\lambda}$$

$$=\lambda^2e^{-\lambda}\sum_{k=2}^{\infty}\frac{\lambda^{k-2}}{(k-2)!}+\lambda e^{-\lambda}\sum_{k=1}^{\infty}\frac{\lambda^{k-1}}{(k-1)!}$$

$$=\lambda^2e^{-\lambda}e^{\lambda}+\lambda e^{-\lambda}e^{\lambda}=\lambda^2+\lambda$$

$$D\xi=E\xi^2-(E\xi)^2=\lambda^2+\lambda-\lambda^2=\lambda$$

普阿松分布的均值与方差相等，都等于参数λ。

例2.5.10 若ξ的均值$E\xi$、方差$D\xi$都存在，求随机变量$\eta=\frac{\xi-E\xi}{\sqrt{D\xi}}$的均值与方差

解 $E\eta=\frac{1}{\sqrt{D\xi}}E(\xi-E\xi)=\frac{1}{\sqrt{D\xi}}(E\xi-E\xi)=0$

$$D\eta=\frac{1}{D\xi}D(\xi-E\xi)=\frac{D\xi}{D\xi}=1$$

随机变量η称为ξ的标准化。

（三）协方差、相关系数

对于两个随机变量，除了需要考虑它们各自的均值与方差外，有时还要考虑它们之间的关系。下面定义的协方差就是反映两个随机变量之间关系的数字特征。

定义2.5.3 如果随机变量ξ，η及$\xi\eta$的均值存在，则称

$$\begin{aligned}cov(\xi,\eta)&=E[(\xi-E\xi)(\eta-E\eta)]\\&=E(\xi\eta)-(E\xi)(E\eta)\end{aligned} \tag{2.5.9}$$

为ξ与η的协方差，而

$$\rho=\frac{cov(\xi,\eta)}{\sqrt{D\xi}\sqrt{D\eta}} \tag{2.5.10}$$

称为ξ与η的相关系数。

相关系数ρ显然是ξ与η的标准化

$$\frac{\xi-E\xi}{\sqrt{D\xi}}与\frac{\eta-E\eta}{\sqrt{D\eta}}$$

的协方差。

在进一步说明相关系数的性质之前，先证明下面的不等式。

许瓦兹（Schwarz）不等式 若对随机变量ξ与η，$E(\xi^2)$与$E(\eta^2)$存在，则$E(\xi\eta)$也存在，且

$$[E(\xi\eta)]^2\leqslant E(\xi^2)\cdot E(\eta^2) \tag{2.5.11}$$

证明 由不等式$|ab|\leqslant\frac{1}{2}(a^2+b^2)$[$\because(|a|-|b|)^2\geqslant0$]可知$E(\xi\eta)$存在。

又因为对任意实数 t，随机变量$(t\xi+\eta)^2$ 非负，所以

$$E[(t\xi+\eta)^2]=t^2E(\xi^2)+2tE(\xi\eta)+E(\eta^2)\geqslant 0$$

因此，

$$[E(\xi\eta)]^2-E(\xi^2)\cdot E(\eta^2)\leqslant 0 \qquad (2.5.12)$$

于是不等式得证。

相关系数具有下面的性质

性质 1　$|\rho|\leqslant 1$

证明　在许瓦兹不等式中，分别用$\xi-E\xi$与$\eta-E\eta$代替ξ与η，便得

$$[E(\xi-E\xi)(\eta-E\eta)]^2\leqslant E(\xi-E\xi)^2\cdot E(\eta-E\eta)^2$$

即　$cov(\xi,\eta)^2\leqslant D\xi\cdot D\eta$，所以

$$|\rho|\leqslant 1$$

性质 2　若ξ，η独立，则$\rho=0$

证明　若ξ，η 独立，由均值性质（2.5.3）知 $E(\xi\eta)=(E\xi)\cdot(E\eta)$，再由（2.5.9），则得 $cov(\xi,\eta)=0$，因而 $\rho=0$。

注意$\rho=0$并不能推出ξ，η独立，若$\rho=0$，我们称ξ，η不相关。不相关不一定独立。下面就是一例。设（ξ，η）的联合分布由下表给出。

η \ 联合分布 \ ξ	-1	0	1	η的分布
-1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$p(\eta=-1)=\frac{3}{8}$
0	$\frac{1}{8}$	0	$\frac{1}{8}$	$p(\eta=0)=\frac{2}{8}$
1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$p(\eta=1)=\frac{3}{8}$
ξ的分布	$p(\xi=-1)=\frac{3}{8}$	$p(\xi=0)=\frac{2}{8}$	$p(\xi=1)=\frac{3}{8}$	

我们来说明 ξ 与 η 的相关系数 $\rho=0$，但 ξ 与 η 不独立。事实上，由于 $E\xi=0$，$E\eta=0$，$E(\xi\eta)=(-1)\frac{1}{4}+0\times\frac{1}{2}+(+1)\frac{1}{4}=0$，

所以，$\mathrm{cov}(\xi,\eta)=0$，从而 $\rho=0$。但，由于 $P(\xi=-1,\eta=1)=\frac{1}{8}$

$\neq P(\xi=-1)\cdot P(\eta=1)=\frac{9}{64}$，所以，$\xi$ 与 η 不独立。

性质3　（1）若 $\eta=a\xi+b$，则 $|\rho|=1$；（2）若 $|\rho|=1$，则 $P(\eta=a\xi+b)=1$。

证明　（1）若 $\eta=a\xi+b$，则

$$
\begin{aligned}
E(\xi-E\xi)(\eta-E\eta)&=E(\xi-E\xi)(a\xi+b-aE\xi-b)\\
&=aE(\xi-E\xi)^2=aD\xi
\end{aligned}
$$

而　$$D\eta=E(\eta-E\eta)^2=E(a\xi+b-aE\xi-b)^2=a^2D\xi$$

所以

$$\rho=\frac{E(\xi-E\xi)(\eta-E\eta)}{\sqrt{D\xi}\sqrt{D\eta}}=\frac{aD\xi}{|a|D\xi}\qquad\therefore|\rho|=1$$

若 $a>0$ 则 $\rho=1$；若 $a<0$，则 $\rho=-1$。

（2）$\zeta\triangleq[\sqrt{D\eta}(\xi-E\xi)-\sqrt{D\xi}(\eta-E\eta)]^2$

则

$$E\zeta=D\xi D\eta-2\sqrt{D\xi}\sqrt{D\eta}E(\xi-E\xi)(\eta-E\eta)+D\xi D\eta$$

若 $\rho=1$，则 $E(\xi-E\xi)(\eta-E\eta)=\sqrt{D\xi}\sqrt{D\eta}$

因此　$E\zeta=2D\xi D\eta-2D\xi D\eta=0$

又由于 $\zeta\geqslant0$，所以，$P(\zeta=0)=1$

即　$$P(\sqrt{D\eta}(\xi-E\xi)-\sqrt{D\xi}(\eta-E\eta)=0)=1$$

但　上式左端 $$=P\left(\frac{\sqrt{D\eta}}{\sqrt{D\xi}}(\xi-E\xi)=\eta-E\eta\right)$$

$$=P\left(\eta=\frac{\sqrt{D\eta}}{\sqrt{D\xi}}(\xi-E\xi)+E\eta\right)$$

$$=P(\eta=a\xi+b)$$

故 $$P(\eta=a\xi+b)=1$$

这里 $$a=\frac{\sqrt{D\eta}}{\sqrt{D\xi}} \qquad b=E\eta-\frac{\sqrt{D\eta}}{\sqrt{D\xi}}E\xi$$

若 $\rho=-1$ 对 $\zeta\triangleq[\sqrt{D\eta}(\xi-E\xi)+\sqrt{D\xi}(\eta-E\eta)]^2$ 同样可证。

（四）矩

上面讲的均值与方差这两个数字特征，是更一般的数字特征——矩的两个特殊情况。矩的概念在第四、五章中将要用到。

定义2.5.4 设 ξ 为一随机变量，若 ξ^k 的均值

$$m_k=E\xi^k \qquad (k=1,\ 2,\ \cdots,)$$

存在，则称 m_k 为 ξ 的 k 阶原点矩。若 $|\xi^k|$ 的均值

$$a_k=E|\xi^k| \qquad (k=1,\ 2,\ \cdots,)$$

存在，则称 a_k 为 ξ 的 k 阶原点绝对矩。

$k=1$ 时，m_1 就是均值。

定义2.5.5 设 ξ 为一随机变量，若 m_1 存在，则 ξ 关于 m_1 的 k 阶矩

$$\mu_k=E(\xi-m_1)^k$$

存在时，称 μ_k 为 ξ 的 k 阶中心矩。若 ξ 的均值 m_1 存在，则当

$$\beta_k=E|\xi-m_1|^k$$

存在时，称 β_k 为 ξ 的 k 阶中心绝对矩。

当 $m_1=0$ 时，则 $\mu_k=m_k$。显然，$\mu_1=0$，$\mu_2=D\xi$。

中心矩与原点矩有下述关系

$$\mu_n=E(\xi-E\xi)^n=\sum_{k=0}^{n}\binom{n}{k}(-E\xi)^{n-k}E\xi^k$$

$$=\sum_{k=0}^{n}\binom{n}{k}(-E\xi)^{n-k}m_k$$

$$=\sum_{k=2}^{n}(-1)^{n-k}\binom{n}{k}m_k m_1^{n-k}+(-1)^{n-1}(n-1)m_1^n$$

我们仅就 n 的头四个数值写出这个关系

$$\mu_0=1,\ \mu_1=0,\ \mu_2=m_2-m_1^2\quad \mu_3=m_3-3m_2m_1+2m_1^3$$
$$\mu_4=m_4-4m_3m_1+6m_2m_1^2-3m_1^4$$

反之，如果已知 m_1，则由中心矩亦可计算原点矩．因为

$$m_n=E\xi^n=E[(\xi-m_1)+m_1]^n=\sum_{k=0}^{n}\binom{n}{k}E(\xi-m_1)^{n-k}m_1^k$$
$$=\sum_{k=0}^{n}\binom{n}{k}\mu_{n-k}m_1^k$$

习　　题

1　设基本事件空间 $\Omega=(\omega_1,\ \omega_2,\ \omega_3,\ \omega_4,\ \omega_5)$，各基本事件的概率为 $P(\{\omega_i\})=p_i$，$i=1,\ 2,\ 3,\ 4,\ 5$，$\sum_{i=1}^{5}p_i=1$．今定义随机变量为 $\xi(\omega_1)=0$．$\xi(\omega_3)=0$，$\xi(\omega_2)=1$，$\xi(\omega_4)=1$，$\xi(\omega_5)=2$．试求事件

$$A_0=(\omega:\xi(\omega)=0),\ A_1=(\omega:\xi(\omega)=1),\ A_3=(\omega:\xi(\omega)=2)$$

及它们的概率 $P(A_0)$，$P(A_1)$，$P(A_2)$．

2　将一枚硬币扔三次，求基本事件空间 Ω，设 ξ 为三次中出现正面的次数，求

$P(\xi=k)\qquad k=0,\ 1,\ 2,\ 3$

$P(\xi\leqslant 1)$

$P(0<\xi<3)$

3　设 ξ 的概率分布为

$$P(\xi=j)=p_j==\frac{c}{2^j},\qquad j=1,\ 2,\ \cdots$$

求 c 值．

4　设 ξ 的概率分布为

$$P(\xi=j)=p_j=\frac{c}{2+j},\qquad j=0,\ 1,\ 2,\ 3$$

求 c 值及下列概率

$P(\xi=3)$，$P(\xi<3)$，$P\{(\xi=2)\cup(\xi=3)\}$

5　已知 n 只电容器中有一只已被击穿，为把这只被击穿了的电容器挑出，我们逐只作检验，以 ξ 表示需作检验的次数，求 ξ 的概率分布。

6　将一枚硬币扔三次，设 ξ 为头两次中出现正面的次数，η 为第3次出现正面的次数，求 (ξ, η) 的联合分布及 ξ，η 的边缘分布。

7　设随机变量 ξ 和 η 相互独立且具有共同的几何分布。令 ζ 为 ξ，η 中的较大者，即

$$\zeta=\max(\xi, \eta)$$

试求 ζ 和 ξ 的联合分布及 ζ 的分布

8　设随机变量 ξ 的概率分布为

$$P(\xi=k)=\frac{1}{2^k}, \quad k=1, 2, 3, \cdots$$

求

（1）$P(\xi$ 为偶数$)$　　（2）$P(\xi\geqslant 5)$

（3）$P(\xi$ 为 3 的倍数$)$

9　假定有 $n=5$ 个工人独立地工作，假定每个工人在一小时内平均有12分钟需要电力。

（1）求在同一时刻有 3 个工人需要电力的概率。

（2）如果最多只能供应 3 个人需要的电力，求超过负荷的概率。

10　若产品的废品率 $p=0.01$。

（1）问需要取多少件产品，才能使一件废品也没有的概率 $\geqslant 0.95$?

（2）问需要取多少件产品，才能使至少出现一件废品的概率 $\geqslant \frac{1}{2}$?

11　随机数字列要有多长才能使序列中数字 7 出现的概率至少是9/10?

12　设中靶的概率是1/5，问10次独立打靶中，至少有两次中靶的概率是多少？

13　在上题中，如果已知至少有一次中靶，问至少两次中靶的条件概率是多少？

14　设两个人都扔 n 次硬币，求硬币出现正面的次数相等的概率。

15　（1）从含有5件次品的50件产品中，随机取出20件，求20件产品中次品数少于3的概率；（2）一大批产品，其次品率为0.10，从中随机取出20件，求20件产品中次品数少于3的概率。

16　若干支枪同时向目标独立地进行射击，设至少要有三枪击中目标，目标才能被击毁，并设每支枪射击的命中率都是0.4，问至少要有多少支枪才能使击毁目标的概率$\geqslant 0.9$？

17　将贝努利试验独立地重复 n 次，假定在每次试验中成功的概率为 p，失败的概率为 $q=1-p$，求"在恰有 r 次成功的条件下，第 i 次试验成功"的概率。

18　某售货员同时出售两包同样的书，每包有 N 本，每次售书他等可能地任选一包，从中取出一本，直到他发现一包已空为止。问这时另一包中尚余 r 本书的概率是多少？

19　设敌机俯冲时被一支步枪击中的概率是0.008，用普阿松定理计算，当250支步枪同时开火时，击中敌机的概率。

20　一无线电仪器由1000个元件组成。每一元件在一年内发生故障的概率是0.001，并且和其他元件的状况无关。用普阿松定理计算，在一年内（1）恰有2个元件发生故障的概率；（2）不少于2个元件发生故障的概率。

21　设一本500页的书中含有500个错处，试计算一页中至少有三个错处的概率。

22　一本书包含 n 页，每一页的印错处平均为 λ，计算至少有一页其中印错之处超过 k 的概率。

23　ξ 服从普阿松分布，若已知$P(\xi=0)=0.4$，求$P(\xi \quad 2)$。

24　已知每天到达某炼油厂的油船数ξ服从参数为$\lambda=2$的普阿松分布，港口一天只能服务三只油船，如果一天中到达的油船超过三只，则超过三只的油船必须转向另一港口。

（1）求在一天中必须有油船转走的概率。

（2）设备增加到多少（即每天能服务多少只船），才能使每天有0.90的船只得到服务。

（3）求每天船到的最大可能数。

（4）求每天平均服务船数。

（5）求每天平均离开的船数

25　在7个时间区间里观察放射性源，每一个时间区间为10秒．假设在每个时间区间里放出粒子数ξ服从参数为5的普阿松分布。求

（1）在每个时间区间里有$\geqslant 4$个粒子放出的概率。

（2）至少有一个区间有$\geqslant 4$个粒子放出的概率。

26　从放射性源放出粒子，假设在一小时内放出的粒子数ξ服从参数$\lambda=4$的普阿松分布，并假设记录放射出的粒子时，有时记漏，记漏的概率为$p=0.1$。求在一小时内记录的粒子数η等于0的概率$P(\eta=0)$。

27　设一只昆虫产r个卵的概率为$\frac{\lambda^r}{r!}e^{-\lambda}$，$\lambda>0$，而卵能发育成虫的概率为$p$。又设每个卵是否发育成虫是彼此独立的，求一只昆虫有k个后代的概率。

28　已知ξ的概率分布为$p(-2)=p(-1)=p(1)=p(2)=\frac{1}{4}$，求$\eta=\xi^2$的概率分布。

29　上题中，求ξ，η的联合分布$p(x, y)$及协方差$\mathrm{cov}(\xi,\eta)$，并问ξ与η是否独立。

30　已知ξ，η相互独立，ξ服从参数为n，p的二项分布，η服从参数为m，p的二项分布，求$\xi+\eta$的概率分布。

31 若ξ的取值为 $x_i=(-1)^{i+1}3^i/i$，而 $P(\xi=x_i)=2/3^i$ $(i=1, 2, \cdots)$ 问 $E(\xi)$ 是否存在？

32 若ξ的概率分布为。

$$\begin{pmatrix} 0, & 1, & 2 \\ \frac{1}{2}, & \frac{3}{8}, & \frac{1}{8} \end{pmatrix}$$

求 $E\xi$，$E\xi^2$，$D\xi$。

33 证明：若ξ的概率分布为

$$\begin{pmatrix} 1, & 2, & 3, & \cdots\cdots \\ p_1, & p_2, & p_3, & \cdots\cdots \end{pmatrix}$$

则 $$E\xi=\sum_{k=1}^{\infty} kp_k=\sum_{k=0}^{\infty} P(\xi>k)$$

34 证明：若$a\leqslant\xi\leqslant b$，则$a\leqslant E\xi\leqslant b$ $\quad D\xi\leqslant\left(\frac{b-a}{2}\right)^2$ 。

35 证明：（1）$D(\xi+\eta)=D\xi+D\eta+2\text{cov}(\xi, \eta)$；（2）$\rho(a\xi+b, c\eta+d)=\rho(\xi, \eta)$，其中 a, b, c, d 为常数，$a>0$，$c>0$。

36 设随机变量ξ与η独立，且 $E\xi=2$，$D\xi=1$，$E\eta=1$，$D\eta=4$，求 $\zeta_1=\xi-2\eta$，$\zeta_2=2\xi-\eta$ 的均值与方差。

37 令ξ为可列重贝努利试验中从第一次试验开始的连贯（结果相同）的长度，求 $E\xi$。

38 已知总机在一分钟内平均接到60次呼唤，试求在接线员离开的30秒钟内，一次呼唤也没有的概率。

39 设三个人生产同一种零件，若三个人生产好品的概率分别为0.95，0.80，0.85，现在他们分别生产了n_1，n_2，n_3件产品，求三个人所生产的$n_1+n_2+n_3$件产品中好品的平均数。

40 设已知(ξ, η)的联合分布为

$$p(1, 5)=0, \ p(1,10)=\frac{1}{2}, \ p(2, 5)=\frac{1}{2}, \ p(2, 10)=0.$$

求 $E\xi$，$E\eta$，$D\xi$，$D\eta$，$\text{cov}(\xi, \eta)$，$\rho(\xi, \eta)$。

41 有一个人想开门，他共有n把钥匙，其中只有一把能打开门。他从n把中随机取一把，如果试验结果开不开，则这把钥匙不

放回，然后再取一把，如此进行下去直到开开门为止，试求所需试验次数 ξ 的均值与方差。

42　N（相当大的数目）个人都去验血，这可以用两种办法进行：（1）每个人都分别去验，这时要求化验 N 次；（2）把 k 个人的血混在一起化验，如果化验结果正常，那么对这 k 个人来说只作一次化验就行了，如果化验结果说明混合血不正常，那么必须对这 k 个人再逐个地分别化验，这时对这 k 个人就需要作 $k+1$ 次化验。假定对所有的人来说，血液不正常的概率都是 p，并且把这些人都看成是统计独立的，问

（1）k 个人的混合血不正常的概率是多少？

（2）按第二方案验血，k 个人需要进行化验的次数 ξ 的均值 $E\xi$ 是多少？

43　假设 A 与 B 是两个事件，且 $P(A)>0$，$P(B)>0$，令 ξ 与 η 是如下定义的随机变量：

$$\xi=\begin{cases}1 & A\text{发生}\\ 0 & A\text{不发生}\end{cases}$$

$$\eta=\begin{cases}1 & B\text{发生}\\ 0 & B\text{不发生}\end{cases}$$

证明由 $\rho_{\xi\eta}=0$ 可推出 ξ，η 独立。

学 习 指 导

（一）本章概要

这一章我们引进了随机变量的概念，并着重讨论了离散型随机变量。概率分布是对随机变量变化规律的完整描述，因此，研究随机变量重要的是研究它的概率分布。离散型随机变量的概率分布可用分布列或概率函数来表示。这一章我们着重研究了一维的情况，同时也介绍了二维的情况，关于二维离散型随机变量要用联合分布

来描述，了解联合分布与边缘分布的关系是很重要的。

关于离散型随机变量的概率分布着重介绍了两种常见的类型，即二项分布与普阿松分布，它们不论在理论上还是在应用上都是非常重要的。

随机变量的数字特征是描述随机变量的某一侧面特征的数值。这一章主要介绍了两个重要的数字特征，即均值与方差。关于两个随机变量的关系，我们引进了协方差、相关系数、独立性的概念及条件分布的概念。此外，关于随机变量函数的分布也是很重要的。

（二）内容说明

1　关于概率函数的性质

离散型随机变量的概率函数有下面两个性质

（1）$P(\xi=x_i)=p(x_i)\geqslant 0 \qquad i=1,\ 2,\ \cdots$

（2）$\sum\limits_{i=1}^{\infty}p(x_i)=1$

定义在ξ的可能取值$\{x_1,\ x_2,\ \cdots\}$上的函数$p(x_i)$，如果满足（1），（2）则它就是一个概率函数，因为它满足概率公理。因此，可用这两个性质来验证某个函数$p(x_i)$是否为概率函数。比如下面写的就不是一个随机变量的分布列

$$\begin{pmatrix} 1, & 2, & 3, & \cdots, & n-1 \\ \frac{1}{n}, & \frac{1}{n}, & \frac{1}{n}, & \cdots, & \frac{1}{n} \end{pmatrix}$$

因为$\frac{1}{n}(n-1)\neq 1$。

有时还可用这两个性质来确定概率函数中的未知参数。比如，

例2.1　求满足下面概率函数的未知数a。

（1）$P(\xi=k)=ae^{-k+2}, \qquad k=0,\ 1,\ 2,\ \cdots$

（2）$P(\xi=k)=\dfrac{a}{3^k\cdot k!}, \qquad k=0,\ 1,\ 2,\ \cdots$

解（1）$\sum\limits_{k=0}^{\infty}ae^{-k+2}=ae^2\sum\limits_{k=0}^{\infty}e^{-k}=ae^2\dfrac{e}{e-1}$

令 $\frac{ae^3}{e-1}=1$ 得

$$a=\frac{e-1}{e^3}$$

（2）$\sum_{k=0}^{\infty}\frac{a}{3^k\cdot k!}=a\sum_{k=0}^{\infty}\frac{(1/3)^k}{k!}=ae^{1/3}$

令 $ae^{1/3}=1$ 得

$$a=e^{-1/3}$$

2 关于二项分布

（1）贝努利试验

一个试验常常是有不止两个结果，但是由于我们关心的只是某一个事件A是否发生，这时，我们对这个试验就只考虑事件A及事件$\overline{A}$，所以，可以把这个试验看做是贝努利试验。比如，掷一颗骰子，如果我们只关心它是否出现1点，则这个试验可看做是只有两个结果：A＝“出1点”和$\overline{A}$＝“不出1点”的贝努利试验。

再比如，从0，1，…，9这10个数码中随机取一个数码，这时可能有10个不同的基本结果，如果我们关心的是它出不出现7，这时，我们也可把这个试验看成是 $P(A)=\frac{1}{10}$的贝努利试验，其中A＝“出现7”，$\overline{A}$＝“不出现7”。正因为这样，所以，贝努利试验序列的模型有广泛的应用。比如，

例2.2 随机数字列要有多长才能使在序列中出现数字7的概率$\geqslant 0.9$？

解 设所需长为n，则在长为n的数字列中至少出现一个7的概率就是在n次独立贝努利试验中A至少发生一次的概率，所以为

$$P(\text{数列中至少出现一个 }7)=1-P(\text{数列中不出现 }7)$$
$$=1-(0.9)^n$$

由于要求$1-(0.9)^n\geqslant 0.9$，即$(0.9)^n\leqslant 0.1$，所以得

$$n\geqslant 22$$

（2）二项分布的应用

只要所考虑的试验可看做是贝努利试验的 n 次独立重复，则 n 次试验中 A 发生的次数 ξ 就服从二项分布，有很多内容不同的问题，实际上都是属于二项分布的概率计算问题. 比如，观察 n 个工人中某一时刻"用电"（A）的工人数；观察 n 口猪中"得病"（A）的头数；观察 n 次射击中"击中"（A）的次数；观察 n 个零件中"废品"（A）的个数等等，只要每次观察 $P(A)$ 都不变，则它们都服从二项分布，因此可用二项分布计算有关概率。下面把例2.2.3再做一详细说明. 请注意下面三种不同的提法.

（a）甲、乙两盒同样的零件，每盒 N 件，一工人用零件时，随机从一盒中取一个，连续地取，求事件 B_1 = "甲盒取空，而乙盒尚余 r 件"的概率.

设 A = "取自甲"，$\overline{A}$ = "取自乙"，且 $P(A)=\frac{1}{2}$，则所求概率即为事件"在 $2N-r$ 次试验中 A 出现 N 次"的概率. 应为

$$P(B_1)=\binom{2N-r}{N}\left(\frac{1}{2}\right)^{N-r}\left(\frac{1}{2}\right)^{N}$$

$$=\binom{2N-r}{N}\left(\frac{1}{2}\right)^{2N-r}$$

（b） 与（a）同样条件，求 B_2 = "第一次发现甲盒已空，而乙盒尚余 r 件"的概率.

第一次发现甲盒已空，这说明共取了 $2N-r+1$ 次，而第 $2N-r+1$ 次是取向甲盒的，所以这时所求概率为

$$P(B_2)=P(B_1)\cdot\frac{1}{2}=\binom{2N-r}{N}\left(\frac{1}{2}\right)^{2N-r+1}$$

（c） 与（a）同样条件，求 B_3 = "第一次发现一盒已空，另一盒尚余 r 件"的概率. 结果应为

$$P(B_3)=2P(B_2)=\binom{2N-r}{N}\left(\frac{1}{2}\right)^{2N-r}$$

（3）二项分布随机变量是 n 个两点分布随机变量的和

设ζ为贝努利试验n次独立重复中事件A发生的次数．$\mu_1, \mu_2, \cdots, \mu_n$分别为在第1次，第2次，…，第$n$次贝努利试验中$A$发生的次数（显然它们都是服从两点分布的），则有

$$\zeta = \mu_1 + \mu_2 + \cdots + \mu_n$$

二项分布随机变量的这一表示很有用．因为μ_i的分布极其简单，所以它就使得有关二项分布的讨论也简单化了．比如．

（a）求二项分布的均值

$$E\zeta = E\mu_1 + E\mu_2 + \cdots + E\mu_n = np$$

（b）求二项分布的方差

$$D\zeta = D\mu_1 + D\mu_2 + \cdots + D\mu_n = npq$$

（c）在第四章将会看到用它来求ζ的矩母函数也是很简单的．

3　关于普阿松分布

在实际中有一些是属于二项分布与普阿松分布结合的问题．这类问题看起来好象比较复杂．其实它只不过是全概率公式的应用．看下面的例子．

例2.3　假设在一小时内从放射源放出的粒子数ξ服从参数为λ的普阿松分布，并假定记录放射出的粒子时有漏记的可能．记对的概率为p，求在一小时内记录的粒子数η的概率分布．

解　求$P(\eta = k)$，$k = 0, 1, 2, \cdots$

类似于导出全概率公式的想法，可写出

$$P(\eta = k) = P(\eta = k, \xi = k) + P(\eta = k, \xi = k+1) + P(\eta = k, \xi = k+2) + \cdots$$

对等式右端各概率用一下乘法公式便得

$$P(\eta = k) = \sum_{r=k}^{\infty} P(\eta = k \mid \xi = r) P(\xi = r)$$

$$= \sum_{r=k}^{\infty} \binom{r}{k} p^k q^{r-k} \cdot \frac{\lambda^r}{r!} e^{-\lambda}$$

$$= e^{-\lambda} \sum_{r=k}^{\infty} \frac{r!}{k!(r-k)!} p^k q^{r-k} \frac{\lambda^r}{r!}$$

$$= e^{-\lambda}\frac{(\lambda p)^k}{k!}\sum_{r=k}^{\infty}\frac{(\lambda q)^{r-k}}{(r-k)!}$$

$$= e^{-\lambda}e^{\lambda q}\frac{(\lambda p)^k}{k!}$$

$$= e^{-\lambda p}\frac{(\lambda p)^k}{k!}$$

可知 η 服从参数为 λp 的普阿松分布。

再举一个与例2.3结果完全一样的例子（计算省略）。

例2.4　已知染色体分裂后的细胞数 ξ 服从参数为 λ 的普阿松分布，而每一个分裂成的染色体不再恢复原状的概率为 p，求所观察到的分裂数 η 的概率分布。

4　关于几何分布

关于几何分布的方差，我们是用矩母函数求的，结果为 $\frac{q}{p^2}$，这样求是为了简单。当然用方差定义也可直接求，只是稍微麻烦一点，下面我们直接求一下。

已知几何分布为 $P(\xi=k)=q^{k-1}p$，$k=1, 2, \cdots$ 由于已知 $E\xi=\frac{1}{p}$，而 $D\xi=E\xi^2-(E\xi)^2$，所以只求 $E\xi^2$ 就行了。

$$E\xi^2=\sum_{k=1}^{\infty}k^2q^{k-1}p=p(1+2^2q+3^2q^2+4^2q^3+\cdots)$$

$$=p(q+2q^2+3q^3+4q^4+\cdots)'$$

$$=p[q(1+2q+3q^2+4q^3+\cdots)]'$$

$$=p\left(\frac{q}{(1-q)^2}\right)' \qquad \text{（参看例2.5.5）}$$

$$=p\frac{(1-q)^2+2(1-q)q}{(1-q)^4}=\frac{p+2q}{p^2}=\frac{1+q}{p^2}$$

所以

$$D\xi=\frac{1+q}{p^2}-\frac{1}{p^2}=\frac{q}{p^2}$$

可以把几何分布进一步推广而得到巴斯卡（Pascal）分布。

在 $P(A)=p$ 的贝努利试验序列中，若以 η 表示第 r 次出现 A 时

试验的次数，则

$$P(\eta=k)=\binom{k-1}{r-1}p^r q^{k-r}, \quad k=r,\ r+1,\ \cdots$$

这是因为$(\eta=k)$表示在第k次试验时出现A，而在前$k-1$次试验中A出现了$r-1$次。在$k-1$次试验中A出现$r-1$次的概率为$\binom{k-1}{r-1}p^{r-1}q^{(k-1)-(r-1)}$ 所以

$$P(\eta=k)=\binom{k-1}{r-1}p^{r-1}q^{k-r}\cdot p=\binom{k-1}{r-1}p^r q^{k-r}$$

这个分布称为巴斯卡分布，当$r=1$时，即得几何分布。

例2.5 袋中放有a个黑球和b个白球，从中有放回地一个一个地随机往外取，求直到出现r个白球为止所需抽取的次数的分布。

解 设ξ为直到出现r个白球为止所需抽取的次数，则

$$P(\xi=k)=\binom{k-1}{r-1}\left(\frac{b}{a+b}\right)^{r-1}\left(\frac{a}{a+b}\right)^{k-r}\left(\frac{b}{a+b}\right)$$

$$=\frac{\binom{k-1}{r-1}b^r a^{k-r}}{(a+b)^k}$$

5 关于用概率不等式求解的问题

在概率论中常有如下提法的问题。

（1）已知一批产品的废品率$p=0.01$，问需要检查多少件产品，才能使一件废品也没有的概率大于或等于0.95？

（2）已知步枪击毁目标的概率$p=0.4$，问需要多少支步枪才能使击毁目标的概率不小于0.9？

（3）已知每月出售某种商品的件数ξ服从参数$\lambda=7$的普阿松分布，问库存多少件才能以0.999的概率满足顾客需要？

关于这一类问题，首先应设所求的未知数为n，比如在（1）中设需要检查的产品数为n，然后根据题意写出计算有关事件的概率的式子，比如在（1）中就是

P（检查n件中一件废品也没有）

$$=P(\zeta=0)=0.99^n$$

其中ζ为n件产品中的废品数．最后再把对概率的要求表示出来，即得不等式

$$0.99^n \geqslant 0.95$$

由这个不等式即可求n，此题得$n\leqslant 5$.

再比如（2），首先设所需步枪数为n，然后根据题意写出计算有关事件的概率的式子．

$$P(n\text{支步枪将目标击毁})$$
$$=P(\zeta\geqslant 1)=1-P(\xi=0)=1-(0.6)^n$$

其中ζ为n支步枪中击中目标的枪数．最后把对概率的要求表示出来，即得不等式

$$1-(0.6)^n\geqslant 0.9 \qquad (0.6)^n\leqslant 0.1$$
$$n\geqslant 5$$

关于（3）亦可按类似步骤来求．

6 关于均值

在均值的定义中要求级数$\sum_{k=1}^{\infty}x_k p(x_k)$绝对收敛，否则我们说均值不存在．下面举两个不存在均值的随机变量的例子．

（1）若随机变量ξ的取值为 $x_k=(-1)^k\cdot\frac{2^k}{k}$，$k=1,2,\cdots$，对应的概率为$p_k=\frac{1}{2^k}$，则由于 $p_k\geqslant 0$；$\sum_{k=1}^{\infty}p_k=1$，因此p_k是概率分布．

$$\sum_{k=1}^{\infty}x_k p_k=\sum_{k=1}^{\infty}(-1)^k\frac{2^k}{k}\cdot\frac{1}{2^k}=\sum_{k=1}^{\infty}\frac{(-1)^k}{k}$$
$$=-\ln 2$$

但是由于

$$\sum_{k=1}^{\infty}|x_k|p_k=\sum_{k=1}^{\infty}\frac{1}{k}=\infty$$

所以 ξ 的均值不存在，不能认为 $-\ln 2$ 是它的均值。

（2）若随机变量 ξ 的取值为 $x_k=\frac{(-1)^{k+1}3^k}{k}$，$p_k=P(\xi=x_k)=\frac{2}{3^k}$，$k=1$，2，…，则 $p_k\geqslant 0$，$\sum_{k=1}^{\infty}p_k=\sum_{k=1}^{\infty}\frac{2}{3^k}=1$，因此 p_k 是概率分布。

$$\sum_{k=1}^{\infty}x_k p_k=\sum_{k=1}^{\infty}\frac{(-1)^{k+1}3^k}{k}\cdot\frac{2}{3^k}$$

$$=2\sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}=2\ln 2$$

但是由于

$$\sum_{k=1}^{\infty}|x_k|p_k=2\sum_{k=1}^{\infty}\frac{1}{k}=\infty$$

所以 $E\xi$ 不存在，不能认为 $2\ln 2$ 是它的均值。

7 关于相关系数

（1）ξ 与 η 的相关系数 ρ 是 ξ 与 η 的标准化

$$\frac{\xi-E\xi}{\sqrt{D\xi}}\text{与}\frac{\eta-E\eta}{\sqrt{D\eta}}$$

的协方差。

根据协方差的定义，$\frac{\xi-E\xi}{\sqrt{D\xi}}$ 与 $\frac{\eta-E\eta}{\sqrt{D\eta}}$ 的协方差为

$$cov\left(\frac{\xi-E\xi}{\sqrt{D\xi}},\ \frac{\eta-E\eta}{\sqrt{D\eta}}\right)$$

$$=E\left[\left(\frac{\xi-E\xi}{\sqrt{D\xi}}-0\right)\left(\frac{\eta-E\eta}{\sqrt{D\eta}}-0\right)\right]$$

$$=E\left(\frac{\xi-E\xi}{\sqrt{D\xi}}\cdot\frac{\eta-E\eta}{\sqrt{D\eta}}\right)=\frac{E(\xi-E\xi)(\eta-E\eta)}{\sqrt{D\xi D\eta}}$$

$$=\frac{cov(\xi,\eta)}{\sqrt{D\xi D\eta}}=\rho$$

（2）关于相关系数性质3的证明

若 $|\rho|=1$，则 $P(\eta=a\xi+b)=1$. 下面着重说明一下这个性质的证明. 先考虑 $\rho=1$情况.

由于 $\sqrt{D\eta}(\xi-E\xi)-\sqrt{D\xi}(\eta-E\eta)=0$ 可改写为 $\eta=a\xi+b$ 的形式，其中

$$a=\frac{\sqrt{D\eta}}{\sqrt{D\xi}}\qquad b=E\eta-\frac{\sqrt{D\eta}}{\sqrt{D\xi}}E\xi$$

所以，要证明ξ与η线性相关的概率为1，只要证明

$$P(\sqrt{D\eta}(\xi-E\xi)-\sqrt{D\xi}(\eta-E\eta)=0)=1$$

即可. 而

$$\sqrt{D\eta}(\xi-E\xi)-\sqrt{D\xi}(\eta-E\eta)=0$$

等价于

$$\zeta\triangleq[\sqrt{D\eta}(\xi-E\xi)-\sqrt{D\xi}(\eta-E\eta)]^2=0$$

所以，只要证明 $P(\zeta=0)=1$即可.

对于不取负值的随机变量ζ来说，$P(\zeta=0)=1$和 $E\zeta=0$是一回事，而根据条件$\rho=1$，证明 $E\zeta=0$是比较容易的. 所以证明的顺序是

令 $\zeta=[\sqrt{D\eta}(\xi-E\xi)-\sqrt{D\xi}(\eta-E\eta)]^2$

根据 $\rho=1$证明

$$E\zeta=0\Rightarrow P(\zeta=0)=1$$

所以得

$$P(\sqrt{D\eta}(\xi-E\xi)-\sqrt{D\xi}(\eta-E\eta)=0)=1$$

$$P\left(\eta=\frac{\sqrt{D\eta}}{\sqrt{D\xi}}\xi+E\eta-\frac{\sqrt{D\eta}}{\sqrt{D\xi}}E\xi\right)=1$$

对于$\rho=-1$时，想法一样，证法类似.

（3）相关系数只反映两个随机变量线性相关的程度.

根据性质3可知当$|\rho|=1$时，两个随机变量以概率1存在线性关系. 当$\rho=0$时，两个随机变量没有线性相关关系，但是，可以有

其它的相关关系，所以两个随机变量不一定独立。当两个随机变量相互独立时，它们之间没有相关关系，当然线性关系也没有，所以自然应当是$\rho=0$。

8 几个常见的离散型随机变量的综合列表

名称	概率函数	定义域及参数	均值	方差	矩母函数
几何分布	$q^{k-1}p$	$k=1,2,\cdots$ $0<p<1$	$\frac{1}{p}$	$\frac{q}{p^2}$	$\frac{pe^t}{1-qe^t}$
二项分布	$\binom{n}{k}p^kq^{n-k}$	$k=0,1,2,\cdots,n$ $0<p<1$	np	npq	$[pe^t+(1-p)]^n$
普阿松分布	$\frac{\lambda^k}{k!}e^{-\lambda}$	$k=0,1,2,\cdots$ $\lambda>0$	λ	λ	$e^{\lambda(e^t-1)}$

* 见第四章

习题解答

1 $A_0=\{\omega_1,\ \omega_3\}$，$A_1=\{\omega_2,\ \omega_4\}$，$A_2=\{\omega_5\}$

$$P(A_0)=p_1+p_3,\ P(A_1)=p_2+p_4,\ P(A_2)=p_5$$

2 Ω由2^3个基本事件所构成。它们是

$\omega_1=\{正正正\}$，$\omega_2=\{正正反\}$，$\omega_3=\{正反正\}$，

$\omega_4=\{正反反\}$，$\omega_5=\{反正正\}$，$\omega_6=\{反正反\}$，

$\omega_7=\{反反正\}$，$\omega_8=\{反反反\}$；

$$P\{\xi=0\}=P\{\omega_8\}=\frac{1}{8}$$

$$P(\xi=1)=P(\omega_4,\ \omega_6,\ \omega_7)=\frac{3}{8}$$

$$P(\xi=2)=P(\omega_2,\ \omega_3,\ \omega_5)=\frac{3}{8}$$

$$P(\xi=3)=P(P\omega_1)=\frac{1}{8}$$

$$P(\xi \leqslant 1)=P\{(\xi=0)\cup(\xi=1)\}$$
$$=P(\xi=0)+P(\xi=1)=\frac{1}{8}+\frac{3}{8}=\frac{1}{2}$$
$$P(0<\xi<3)=P\{(\xi=1)\cup(\xi=2)\}$$
$$=P(\xi=1)+P(\xi=2)=\frac{3}{8}+\frac{3}{8}=\frac{3}{4}$$

3　由于$\sum_{i=1}^{\infty} p_i=1 \Rightarrow c\sum_{i=1}^{\infty}\frac{1}{2^i}=1$，

但　$\sum_{i=1}^{\infty}\frac{1}{2^i}=\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\cdots+\frac{1}{2^n}+\cdots=1$，所以 $c=1$。

4　由于　$\sum_{i=0}^{3} p_i=1 \Rightarrow c\sum_{i=0}^{3}\frac{1}{2+i}=1$

$$\Rightarrow c=\frac{1}{\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}}=\frac{60}{77}$$

由此，

$$P(\xi=3)=p_3=\frac{60}{77}\cdot\frac{1}{2+3}=\frac{12}{77}$$
$$P(\xi<3)=1-P(\xi=3)=1-\frac{12}{77}=\frac{65}{77}$$
$$P\{(\xi=2)\cup(\xi=3)\}=P(\xi=2)+P(\xi=3)$$
$$=\frac{60}{77}\left[\frac{1}{2+2}+\frac{1}{2+3}\right]=\frac{27}{77}$$

5　以ξ表示需作检验的次数，ξ的可能取值为1，2，…，$n-1$，这是因为不论第$n-1$次挑出的是被击穿的还是未被击穿的，都可把被击穿的找出来，不用再作第n次检验了。

$A_i \triangleq$“第i次检验挑出未被击穿的电容器”

现在求$P(\xi=k)$，$k=1，2，\cdots，n-1$

当$k=1$时　$P(\xi=1)=P(\overline{A}_1)=\frac{1}{n}$

当$k=2，\cdots，n-2$时，

$$P(\xi=k)=P(A_1A_2\cdots A_{k-1}\overline{A}_k)$$
$$=P(A_1)P(A_2|A_1)\cdots P(\overline{A}_k|A_1A_2\cdots A_{k-1})$$
$$=\frac{n-1}{n}\cdot\frac{n-2}{n-1}\cdots\cdots\frac{n-(k-1)}{n-(k-2)}\cdot\frac{1}{n-(k-1)}=\frac{1}{n}$$

当 $k=n-1$ 时，

$$P(\xi=n-1)=P(A_1A_2\cdots A_{n-2})$$

这是因为当检验进行到第 $n-1$ 次时，可能出现两种情况：或者前 $n-2$ 次挑出的均是未被击穿的，而第 $n-1$ 次挑出击穿了的，或者这 $n-1$ 次挑出的均是未被击穿的，这时剩下的一支肯定是击穿了的，就不用再检验了．因此，

$$(\xi=n-1)=A_1A_2\cdots A_{n-2}\overline{A}_{n-1}\cup A_1A_2\cdots A_{n-2}A_{n-1}$$
$$=A_1A_2\cdots A_{n-2}(\overline{A}_{n-1}\cup A_{n-1})$$
$$=A_1A_2\cdots A_{n-2}$$

所以，

$$P(\xi=n-1)=P(A_1A_2\cdots A_{n-2})$$
$$=\frac{n-1}{n}\cdot\frac{n-2}{n-1}\cdots\cdots\frac{n-(n-2)}{n-(n-3)}=\frac{2}{n}$$

于是得 ξ 的分布列为

$$\begin{pmatrix}1, & 2, & \cdots, & n-2, & n-1\\ \frac{1}{n}, & \frac{1}{n}, & \cdots, & \frac{1}{n}, & \frac{2}{n}\end{pmatrix}$$

6　设 ξ 为头两次出现正面的次数，η 为第3次中出现正面的次数，ξ 的可能取值为0，1，2，η 的可能取值为0，1。则

$$P(\xi=0)=\frac{1}{4}\quad P(\xi=1)=\frac{2}{4}\quad P(\xi=2)=\frac{1}{4}$$

$$P(\eta=0)=\frac{1}{2}\quad P(\eta=1)=\frac{1}{2}$$

简略写为

$$P(\xi=k)=\frac{\binom{2}{k}}{4},\ k=0,\ 1,\ 2$$

$$P(\eta=l)=\frac{1}{2}\qquad l=0,\ 1$$

由于ξ与η独立，所以

$$P(\xi=k,\ \eta=l)=P(\xi=k)\cdot P(\eta=l)=\frac{\binom{2}{k}}{4}\cdot\frac{1}{2}=\frac{1}{8}\binom{2}{k}$$

所得分布可列表如下。

P η \ ξ	0	1	ξ的边缘分布
0	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$
1	$\frac{2}{8}$	$\frac{2}{8}$	$\frac{1}{2}$
2	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{4}$
η的边缘分布	$\frac{1}{2}$	$\frac{1}{2}$	

7　先求（ξ，ζ）的联合分布（注意ξ，ζ不独立）。

$$P(\xi=j,\ \zeta=i)=\begin{cases}P(V) & i<j\\ P(\xi=j,\ \eta=i) & i>j\\ P(\xi=j,\ \eta\leqslant i) & i=j\end{cases}$$

由于ξ、η独立，所以，

$$\begin{aligned}P(\xi=j,\ \eta=i)&=P(\xi=j)\cdot P(\eta=i)\\&=q^{j-1}p\cdot q^{i-1}\cdot p=q^{i+j-2}p^2\end{aligned}$$

$$\begin{aligned}P(\xi=i,\ \eta\leqslant i)&=P(\xi=i,\ \bigcup_{k=1}^{i}(\eta=k))=\sum_{k=1}^{i}P(\xi=i,\ \eta=k)\\&=\sum_{k=1}^{i}P(\xi=i)\cdot P(\eta=k)\\&=\sum_{k=1}^{i}q^{i-1}p\cdot q^{k-1}p=p^2q^{i-1}\sum_{k=1}^{i}q^{k-1}\\&=p^2q^{i-1}\frac{1-q^i}{1-q}=q^{i-1}p(1-q^i)\end{aligned}$$

所以，ξ,ζ 的联合分布为

$$P(\xi=j,\zeta=i)=\begin{cases}0\quad, & i<j\\ q^{i+j-2}p^2, & i>j\\ q^{i-1}p(1-q^i), & i=j\end{cases}$$

ζ 的边缘分布为

$$P(\zeta=i)=P\{\bigcup_{j=1}^{\infty}(\xi=j),\ (\zeta=i)\}=\sum_{j=1}^{\infty}P(\xi=j,\ \zeta=i)$$

$$=\sum_{j<i}P(\xi=j,\zeta=i)+P(\xi=i,\zeta=i)$$

$$+\sum_{j>i}P(\xi=j,\ \zeta=i)$$

$$=\sum_{j<i}P(\xi=j,\zeta=i)+P(\xi=i,\zeta=i)$$

$$=\sum_{j=1}^{i-1}q^{i+j-2}p^2+q^{i-1}p(1-q^i)$$

$$=p^2q^{i-2}\sum_{j=1}^{i-1}q^j+q^{i-1}p(1-q^i)$$

$$=p^2q^{i-2}\frac{q-q^i}{1-q}+q^{i-1}p(1-q^i)$$

$$\left(\sum_{j=1}^{i-1}q^j=\frac{q-q^i}{1-q}\right)$$

$$=q^{i-1}p(1-q^{i-1})+q^{i-1}p(1-q^i)$$

$$=2q^{i-1}p-q^{2(i-1)}p-q^{2i-1}p$$

8　(1)　$P(\xi\text{为偶数})=\sum_{i=1}^{\infty}\frac{1}{2^{2i}}=\frac{\frac{1}{4}}{1-\frac{1}{4}}=\frac{1}{3}$

(2)　$P(\xi\geqslant5)=\sum_{i=5}^{\infty}\frac{1}{2^i}=\frac{\frac{1}{2^5}}{1-\frac{1}{2}}=\frac{1}{2^4}=\frac{1}{16}$

（3） $P(\xi\text{为3的倍数})=\sum_{i=1}^{\infty}\frac{1}{2^{3i}}=\frac{\frac{1}{2^3}}{1-\frac{1}{8}}=\frac{1}{7}$.

9　每个工人在某个时刻需要电力的概率为 $p=\frac{12}{60}=\frac{1}{5}$.

设某个时刻需要电力的人数为 ξ，则

（1）在同一时刻有3个工人需要电力的概率为

$$P(\xi=3)=\binom{5}{3}\left(\frac{1}{5}\right)^3\left(\frac{4}{5}\right)^2=0.0512,$$

（2）超负荷的概率即在同一时刻有≥4个工人需用电力的概率为

$$P(\xi\geqslant 4)=\binom{5}{4}\left(\frac{1}{5}\right)^4\left(\frac{4}{5}\right)+\left(\frac{1}{5}\right)^5=0.00672$$

10　设 n 为所需要取的产品数，ξ 为 n 件产品中的废品数，则

（1） n 件中有0件废品的概率为

$$P(\xi=0)=(0.99)^n$$

要求它≥0.95，即 $(0.99)^n\geqslant 0.95$，得

$$n\leqslant 5$$

（2） n 件中至少出现一件废品的概率为

$$P(\xi\geqslant 1)=1-P(\xi=0)=1-(0.99)^n$$

要求它 $\geqslant\frac{1}{2}$，即 $1-(0.99)^n\geqslant\frac{1}{2}$　$(0.99)^n\leqslant\frac{1}{2}$，得

$$n\geqslant 69$$

11　见第二章学习指导的2

12　设 ξ 为10次打靶中中靶的次数，则所求概率为

$$P(\xi\geqslant 2)=1-P(\xi=0)-P(\xi=1)$$
$$=1-\left(\frac{4}{5}\right)^{10}-\binom{10}{1}\left(\frac{1}{5}\right)\left(\frac{4}{5}\right)^9$$
$$=0.6242$$

13　所求概率为 $P(\xi\geqslant 2|\xi\geqslant 1)=\frac{P(\xi\geqslant 2,\xi\geqslant 1)}{P(\xi\geqslant 1)}$

$$= \frac{1 - P(\zeta = 0) - P(\zeta = 1)}{1 - P(\zeta = 0)}$$

$$= \frac{0.6242}{0.8926} = 0.6993$$

14　设 ξ，η 分别为甲、乙二人掷 n 次硬币中正面出现的次数。则所求概率为

$$P(\xi = \eta) = \sum_{k=0}^{n} P(\xi = k, \eta = k) = \sum_{k=0}^{n} P(\xi = k)P(\eta = k)$$

$$= \sum_{k=0}^{n} \binom{n}{k}\left(\frac{1}{2}\right)^n \cdot \binom{n}{k}\left(\frac{1}{2}\right)^n$$

$$= \frac{1}{2^{2n}} \sum_{k=0}^{n} \binom{n}{k}^2 = \frac{1}{2^{2n}} \binom{2n}{n}$$

恒等式 $\binom{n}{0}^2 + \binom{n}{1}^2 + \cdots + \binom{n}{k}^2 + \cdots + \binom{n}{n}^2 = \binom{2n}{n}$ 的证明如下：

若一个罐子中装有 n 个白球，n 个黑球，从中任取 n 个球，则这 n 个球中恰有 k 个白球的概率为

$$P_k = \frac{\binom{n}{k} \cdot \binom{n}{n-k}}{\binom{2n}{n}} = \frac{\binom{n}{k} \cdot \binom{n}{k}}{\binom{2n}{n}} \quad \left(\because \binom{n}{n-k} = \binom{n}{k}\right)$$

但恰有 0 个，恰有 1 个，…，恰有 k 个，…，恰有 n 个白球这 $n-1$ 个互斥事件的并是必然事件，由加法定理得

$$1 = P\{U\} = P_0 + P_1 + \cdots + P_k + \cdots + P_n$$

$$= \frac{1}{\binom{2n}{n}} \left[\binom{n}{0}^2 + \binom{n}{1}^2 + \cdots + \binom{n}{n}^2\right]$$

两端同乘 $\binom{2n}{n}$ 得

$$\binom{2n}{n} = \binom{n}{0}^2 + \binom{n}{1}^2 + \cdots + \binom{n}{k}^2 + \cdots + \binom{n}{n}^2$$

15　（1）设 ξ 为次品数，$P(\xi<3)=\sum_{k=0}^{2}P(\xi=k)$

$$=\sum_{k=0}^{2}\frac{\binom{5}{k}\binom{45}{20-k}}{\binom{50}{20}}=0.6900$$

（2）$P(\xi<3)=\sum_{k=0}^{2}P(\xi=k)$

$$=\sum_{k=0}^{2}\binom{20}{k}(0.1)^{k}(0.9)^{20-k}$$

$$=1-\sum_{k\geq 3}\binom{20}{k}(0.1)^{k}(0.9)^{20-k}$$

$=1-0.32307$　（查二项分布表）

$=0.67693$

16　n 支枪中至少有三支枪击中目标，从而击毁目标的概率为

$$P=\sum_{k=3}^{n}\binom{n}{k}(0.4)^{k}(0.6)^{n-k}$$

欲使其$\geqslant 0.9$，查常用数理统计用表得 $n\geqslant 15$。

17　$A_i\triangleq$ "第 i 次试验成功"，

$B\triangleq$ "n 次试验中恰有 r 次成功"

则所求概率为

$$P(A_i|B)=\frac{P(A_iB)}{P(B)}=\frac{p\cdot\binom{n-1}{r-1}p^{r-1}q^{(n-1)-(r-1)}}{\binom{n}{r}p^{r}q^{n-r}}$$

$$=\frac{\binom{n-1}{r-1}p^{r}q^{n-r}}{\binom{n}{r}p^{r}q^{n-r}}=\frac{r}{n}$$

18　（Banach 问题）

售货员售书直到他发现一包空，另一包尚余 r 本，必作了 $2N-r+1$ 次试验．他在最后一次售书时发现左包空，而右包尚余 r 本的概率为

$$\binom{2N-r}{N}\left(\frac{1}{2}\right)^{N-r}\cdot\left(\frac{1}{2}\right)^{N}\cdot\frac{1}{2}$$

同样的分析完全适用于右包空，而左包尚余 r 本书的情况，故所求的概率为

$$2\binom{2N-r}{N}\left(\frac{1}{2}\right)^{N-r}\left(\frac{1}{2}\right)^{N}\cdot\frac{1}{2}=\binom{2N-r}{N}\left(\frac{1}{2}\right)^{2N-r}$$

19　假定只要有一支步枪击中敌机即击落．设250支步枪中击中敌机的步枪数为 ξ，则所求概率为

$$P(\xi\geqslant 1)=1-P(\xi=0)=1-(0.992)^{250}\doteq 0.8657$$

用普阿松定理作近似计算，这时 $\lambda\doteq 250\times 0.008=2$

则

$$P(\xi\geqslant 1)=1-P(\xi=0)\doteq 1-\frac{2^0}{0!}e^{-2}$$

$$=1-0.1353=0.8647$$

20　设 ξ 为1000个元件中一年内发生故障的个数，$\lambda\doteq 1000\times 0.001=1$．则所求概率为

（1）　$P(\xi=2)\doteq\frac{1^2}{2!}e^{-1}=\frac{1}{2}e^{-1}=0.18394$

（2）　$P(\xi\geqslant 2)\doteq\sum\limits_{k=2}^{\infty}\frac{1^k}{k!}e^{-1}=0.26424$　　（查表）

21　$n=500$　　$p=\frac{1}{500}$　　$np=1$　　所以 $\lambda\doteq 1$

设一页中出现错误的个数为 ξ，则所求概率为

$$P(\xi\geqslant 3)\doteq\sum_{k=3}^{\infty}\frac{1^k}{k!}e^{-1}=0.08030\qquad\text{（查表）}$$

22　设一页中印错的个数为 ξ，则有

$$P(\xi\leqslant k)=\sum_{i=0}^{k}\frac{\lambda^i}{i!}e^{-\lambda}$$

于是

P（n 页中每一页的印错个数均 $\leqslant k$）

$$=[P(\xi\leqslant k)]^n$$

因此，至少有一页的印错数超过 k 的概率为

$$1-[P(\xi\leqslant k)]^n$$

23　$P(\xi=0)=0.4\ \Rightarrow\ e^{-\lambda}=0.4\ \Rightarrow\ \lambda=0.9163$

$$P(\xi>2)=1-P(\xi=0)-P(\xi=1)-P(\xi=2)$$

$$=1-e^{-\lambda}-\lambda e^{-\lambda}-\frac{\lambda^2}{2}e^{-\lambda}=1-e^{-\lambda}\left[1+\lambda+\frac{\lambda^2}{2}\right]$$

$$=0.06556$$

24　（1）在一天必须有油船转走的概率为

$$P(\xi\geqslant4)=\sum_{k=4}^{\infty}\frac{2^k}{k!}e^{-2}=0.14288\qquad（查表）$$

（2）设每天能服务 n 只船，则欲使

$$\sum_{k=0}^{n}P(\xi=k)>0.90，即\sum_{k=n+1}^{\infty}P(\xi=k)\leqslant0.1$$

查表可知 $n+1\geqslant5$，所以得 $n\geqslant4$

（3）$P(\xi=1)=P(\xi=2)=0.27067=\max$　　（查表）

所以，最大可能来到的船数为 1 或 2．

（4）平均服务的船数为

$$1\times0.27067+2\times0.27067+3\times0.32332=1.78197$$

（5）离开的船数为来到船数减 3．比如离开一只船，就意味着来到 4 只船．所以平均离开的船数为

$$1\times0.09022+2\times0.03609+3\times0.01203+4\times0.00343$$
$$+5\times0.00085+6\times0.00019+7\times0.00003+\cdots\doteq0.21781$$

25　已知一个长为10秒的时间区间里，放出的粒子数 ξ 服从参数为 5 的普阿松分布，所以，

（1）$P(\xi\geqslant4)=\sum\limits_{k\geqslant4}\frac{5^k}{k!}e^{-5}=0.73497$

所以，在 7 个时间区间里每个区间中放出的粒子数 ξ 均 $\geqslant4$ 的概率为

$$[P(\xi\geqslant 4)]^7=(0.73497)^7=0.11585$$

（2）至少有一个区间有≥4个粒子放出的概率为

$$1-[P(\xi<4)]^7=1-(1-0.73497)^7=1-0.26503^7$$

$$=0.99991.$$

26　在一小时内记录的粒子数η等于0的概率为

$$P(\eta=0)=P\{(\eta=0),\bigcup_{k=0}^{\infty}(\xi=k)\}=P\{\bigcup_{k=0}^{\infty}(\eta=0,\xi=k)\}$$

$$=\sum_{k=0}^{\infty}P(\eta=0,\ \xi=k)=\sum_{k=0}^{\infty}P(\xi=k)\cdot P(\eta=0|\xi=k)$$

$$=\sum_{k=1}^{\infty}\frac{4^k}{k!}e^{-4}(0.1)^k+e^{-4}\times 1\quad(\because P(\eta=0|\xi=0)=1)$$

$$=e^{-4}\sum_{k=0}^{\infty}\frac{(0.4)^k}{k!}=e^{-4}\cdot e^{0.4}$$

$$=e^{-3.6}$$

$$=0.02732$$

27　设昆虫产卵数为ξ，一只昆虫的后代数为η，则一只昆虫有k个后代的概率为

$$P(\eta=k)=\sum_{r=k}^{\infty}P(\xi=r)P(\eta=k|\xi=r)$$

$$=\sum_{r=k}^{\infty}\frac{\lambda^r}{r!}e^{-\lambda}\binom{r}{k}p^kq^{r-k}$$

$$=\sum_{r=k}^{\infty}\frac{\lambda^r}{r!}e^{-\lambda}\cdot\frac{r!}{k!(r-k)!}p^kq^{r-k}$$

$$=e^{-\lambda}\frac{(\lambda p)^k}{k!}\sum_{r=k}^{\infty}\frac{(\lambda q)^{r-k}}{(r-k)!}$$

$$=e^{-\lambda}\frac{(\lambda p)^k}{k!}e^{\lambda(1-p)}$$

$$=\frac{(\lambda p)^k}{k!}e^{-\lambda p}$$

可知 η 服从参数为 λp 的普阿松分布。

28　已知 ξ 的分布列为

$$\begin{pmatrix} -2, & -1, & 1, & 2 \\ \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4}, & \frac{1}{4} \end{pmatrix}$$

$\eta=\xi^2$ 的可能取值为 1，4，

$$P(\eta=1)=P(\xi=-1)+P(\xi=1)=\frac{1}{2}$$

$$P(\eta=4)=P(\xi=-2)+P(\xi=2)=\frac{1}{2}$$

所以 η 的分布列为

$$\begin{pmatrix} 1, & 4 \\ \frac{1}{2}, & \frac{1}{2} \end{pmatrix}$$

29　(ξ,η) 的可能取值为

$$(-1,\ 1),\ (1,\ 1),\ (-2,\ 4)\ \ (2,\ 4)$$

$$p(-1,1)=p(1,1)=p(-2,4)=p(2,4)=\frac{1}{4}$$

$$cov(\xi,\eta)=E(\xi-E\xi)(\eta-E\eta)=E(\xi\eta)-(E\xi)(E\eta)$$

$E\xi=0$，$\xi\eta$ 可能取值为 -1×1，1×1，-2×4，2×4，其余情况，概率为 0。所以

$$E(\xi\eta)=-1\times\frac{1}{4}+1\times\frac{1}{4}-8\times\frac{1}{4}+8\times\frac{1}{4}=0$$

所以，　$cov(\xi,\ \eta)=0$

ξ，η 不独立，因为 $P(\xi=k,\ \eta=l)\neq P(\xi=k)\cdot P(\eta=l)$

比如，

$$P(-2,4)=\frac{1}{4}\neq P(\xi=-2)\cdot P(\eta=4)=\frac{1}{4}\cdot\frac{1}{2}=\frac{1}{8}$$

30　可用两种方法求。

（1）$\xi+\eta$ 可能取值为 0，1，2，…，$n+m$。

$$P(\xi+\eta=l)=\sum_{k=0}^{l}P(\xi=k,\eta=l-k)=\sum_{k=0}^{l}$$

$$P(\xi=k)\cdot P(\eta=l-k)=\sum_{k=0}^{l}\binom{n}{k}p^{k}q^{n-k}\cdot\binom{m}{l-k}p^{l-k}q^{m-l+k}$$

$$=\sum_{k=0}^{l}\binom{n}{k}\binom{m}{l-k}p^{l}q^{m+n-l}$$

$$=\binom{n+m}{l}p^{l}q^{m+n-l}$$

（2）设 $\mu_1, \mu_2, \cdots, \mu_n, \mu_1', \mu_2', \cdots, \mu_m'$ 均服从两点分布（参数为 p），则

$$\xi=\mu_1+\mu_2+\cdots+\mu_n,\quad \eta=\mu_1'+\mu_2'+\cdots+\mu_m'$$

$$\xi+\eta=\mu_1+\mu_2+\cdots+\mu_n+\mu_1'+\mu_2'+\cdots+\mu_m'$$

所以，$\xi+\eta$ 服从参数为 $(n+m, p)$ 的二项分布。

31 见第二章学习指导的 7

32 $E\xi=1\cdot\frac{3}{8}+2\cdot\frac{1}{8}=\frac{5}{8}$

$$E\xi^2=1\cdot\frac{3}{8}+4\cdot\frac{1}{8}=\frac{7}{8}$$

$$D\xi=E\xi^2-(E\xi)^2=\frac{56}{64}-\frac{25}{64}=\frac{31}{64}$$

33 $\sum_{k=0}^{\infty}P(\xi>k)=\sum_{k=1}^{\infty}P(\xi\geqslant k)=\sum_{k=1}^{\infty}\sum_{i=k}^{\infty}P(\xi=i)$

$$=\sum_{i=1}^{\infty}[P(\xi=k)+P(\xi=k+1)+\cdots]$$

$$=p_1+p_2+p_3+p_4+p_5+\cdots+p_2+p_3+p_4+p_5+\cdots$$

$$+p_3+p_4+p_5+\cdots$$

$$=p_1+2p_2+3p_3+4p_4+\cdots=E\xi$$

34 $E\xi=\sum_{i=1}^{\infty}x_ip_i$

由于对一切 i 有 $a\leqslant x_i\leqslant b$

所以 $a\sum_{i=1}^{\infty}p_i\leqslant\sum_{i=1}^{\infty}x_ip_i\leqslant b\sum_{i=1}^{\infty}p_i$

以而得

$a \leqslant E\xi \leqslant b$

由方差的性质3　$D\xi < E(\xi - c)^2$ 其中 $c \neq E\xi$，可知

$$D\xi \leqslant E\left(\xi - \frac{a+b}{2}\right)^2 \leqslant \left(b - \frac{a+b}{2}\right)^2 = \left(\frac{b-a}{2}\right)^2$$

35　（1）$D(\xi+\eta) = E(\xi+\eta)^2 - [E(\xi+\eta)]^2$

$$= E(\xi^2 + 2\xi\eta + \eta^2) - [(E\xi)^2 + (E\eta)^2 + 2(E\xi)(E\eta)]$$

$$= [E\xi^2 - (E\xi)^2] + [E\eta^2 - (E\eta)^2] + 2[E\xi\eta - (E\xi)(E\eta)]$$

$$= D\xi + D\eta + 2\mathrm{cov}(\xi, \eta)$$

（2）$\rho(a\xi + b,\ c\eta + d)$

$$= \frac{E[a\xi + b - E(a\xi + b)][c\eta + d - E(c\eta + d)]}{\sqrt{D(a\xi + b) \cdot D(c\eta + d)}}$$

$$= \frac{E[a(\xi - E\xi) \cdot c(\eta - E\eta)]}{\sqrt{a^2 D\xi \cdot c^2 D\eta}}$$

$$= \frac{E(\xi - E\xi)(\eta - E\eta)}{\sqrt{D\xi D\eta}} = \rho(\xi, \eta)$$

36　$E\zeta_1 = E(\xi - 2\eta) = E\xi - 2E\eta = 2 - 2 \times 1 = 0$

$E\zeta_2 = E(2\xi - \eta) = 2E\xi - E\eta = 2 \times 2 - 1 = 3$

$D\zeta_1 = D(\xi - 2\eta) = D\xi + 4D\eta = 1 + 4 \times 4 = 17$

$D\zeta_2 = D(2\xi - \eta) = 4D\xi + D\eta = 4 \times 1 + 4 = 8$

37　见第二章学习指导的4（$P(\xi = k) = pq^k + qp^k$，$E\xi = \frac{q}{p} + \frac{p}{q}$）

38　以 ξ 表示在30秒内接到的呼唤数，则 ξ 服从普阿松分布，参数 $\lambda = 30$。一次呼唤也没有的概率为

$$P(\xi = 0) = \frac{30^0}{0!} e^{-30} = e^{-30}$$

$$= 9.35762 \times 10^{-14}$$

39　三个人所生产的 $n_1 + n_2 + n_3$ 件产品中好品的平均数为

$$0.95n_1 + 0.80n_2 + 0.85n_3$$

40 先求ξ及η的边缘分布

$$P(\xi=1)=\frac{1}{2}\qquad P(\xi=2)=\frac{1}{2}$$

$$P(\eta=5)=\frac{1}{2}\qquad P(\eta=10)=\frac{1}{2}$$

从而 $$E\xi=1\times\frac{1}{2}+2\times\frac{1}{2}=\frac{3}{2}$$

$$E\eta=5\times\frac{1}{2}+10\times\frac{1}{2}=\frac{15}{2}$$

$$D\xi=E\xi^2-(E\xi)^2=1\times\frac{1}{2}+4\times\frac{1}{2}-\frac{9}{4}=\frac{10-9}{4}=\frac{1}{4}$$

$$D\eta=E\eta^2-(E\eta)^2=5^2\times\frac{1}{2}+10^2\times\frac{1}{2}-\frac{15^2}{4}$$

$$=\frac{25+100}{2}-\frac{225}{4}=\frac{25}{4}$$

$$cov(\xi,\eta)=E(\xi\eta)-(E\xi)(E\eta)$$

$$=1\times10\times\frac{1}{2}+2\times5\times\frac{1}{2}-\frac{3}{2}\times\frac{15}{2}$$

$$=\frac{20}{2}-\frac{45}{4}=\frac{-5}{4}$$

$$\rho=\frac{-5/4}{\sqrt{\frac{1}{4}\times\frac{25}{4}}}=-1$$

41 先求ξ的分布$P(\xi=k)$，$(\xi=k)$意味着前$k-1$次取出的钥匙均打不开门，而第k次取出开门的钥匙.

$$P(\xi=k)=\frac{n-1}{n}\cdot\frac{n-2}{n-1}\cdot\cdots\cdot\frac{n-(k-1)}{n-k+2}\cdot\frac{1}{n-k+1}=\frac{1}{n}$$

$$E\xi=\sum_{k=1}^{n}k\cdot\frac{1}{n}=\frac{1}{n}\sum_{k=1}^{n}k=\frac{1}{n}\cdot\frac{1}{2}n(n+1)=\frac{n+1}{2}$$

$$E\xi^2=\sum_{k=1}^{n}k^2\cdot\frac{1}{n}=\frac{1}{n}\sum_{k=1}^{n}k^2=\frac{1}{6}(n+1)(2n+1)$$

$$\begin{aligned}D\xi&=E\xi^2-(E\xi)^2\\&=\frac{1}{6}(n+1)(2n+1)-\frac{1}{4}(n+1)^2\\&=\frac{1}{12}(n+1)(n-1)\end{aligned}$$

42　（1）k 个人的混合血不正常的概率是

$$1-(1-p)^k$$

（2）ξ 的分布列为

$$\begin{pmatrix}1, & k+1\\(1-p)^k, & 1-(1-p)^k\end{pmatrix}$$

$$\begin{aligned}E\xi&=(1-p)^k+(k+1)[1-(1-p)^k]\\&=(1-p)^k+(k+1)-k(1-p)^k-(1-p)^k\\&=k+1-k(1-p)^k=k+1-kq^k\end{aligned}$$

这是对一组 k 个人所求出的平均化验次数. N 个人，分成 $\frac{N}{k}$ 组，所以总共应该化验的次数的平均数应为

$$\frac{N}{k}[k+1-kq^k]$$

43　$\rho_{\xi\eta}=0\Longleftrightarrow\text{cov}(\xi,\eta)=0\Longleftrightarrow E(\xi\eta)=(E\xi)(E\eta)$

$\Longleftrightarrow 1\times P[(\xi=1)(\eta=1)]=1\times P(\xi=1)\cdot 1\times P(\eta=1)$

$\Longleftrightarrow P[(\xi=1)(\eta=1)]=P(\xi=1)P(\eta=1)$

$\Longleftrightarrow P(A\cap B)=P(A)\cdot P(B)\Longleftrightarrow A$，$B$ 独立

A，B 独立 $\Rightarrow\overline{A}$，B 独立 $\Rightarrow A$，$\overline{B}$ 独立 $\Rightarrow\overline{A}$，$\overline{B}$ 独立

所以

$$P(\xi=0,\ \eta=1)=P(\xi=0)P(\eta=1)$$

$$P(\xi=1,\ \eta=0)=P(\xi=1)P(\eta=0)$$

$$P(\xi=0,\eta=0)=P(\xi=0)P(\eta=0)$$

因此，ξ，η 独立.

第三章　连续型随机变量

这一章我们讲连续型随机变量．什么是连续型随机变量呢？直观地讲，就是这种随机变量的取值不是离散的一些点子，而是某些区间或整个数轴．比如：零件的尺寸，农作物的产量，水库的水位高度等等．

既然连续型随机变量的取值不可列，所以其概率规律性就不能用分布列来刻划．那么，怎样来研究连续型随机变量的概率规律性呢？这就引出了所谓“分布密度”的概念．

§3.1　分布密度和分布函数

（一）分布密度

设ξ是一连续型随机变量，我们的目的是设法掌握其取值的概率规律性．既然我们不能一点一点地去研究它的取值情况，那么，是否可以考虑它取值某个范围的可能性大小的问题呢？比如考虑ξ取值某个区间$(a_1,a_2]$的可能性，即概率$P(a_1<\xi\leqslant a_2)$多大．如果对一切实数$a_1<a_2$，上述概率都知道了，当然ξ的概率规律也就清楚了．所以问题的关键是能否把上面形状的概率都求出来．

我们分两种情况来分析：

1　均匀的情况．先考虑一种简单情况：即ξ只取值于某个区间$(a,b]$，而且其取值的概率在这个区间上是均匀的．所谓概率均匀分布在$(a,b]$上，由§1.4可知，就是随机变量ξ取值于$(a,b]$中任一小区间$(a_1,a_2]$的概率大小与这个小区间的区间长度a_2-a_1成正

比，即

$$P(\alpha_1<\xi\leq\alpha_2)=p(\alpha_2-\alpha_1)$$

式中 p 是比例常数．因为 ξ 只能取值于 $(a,b]$ 中，故

$$P(a<\xi\leq b)=p(b-a)=1$$

由此推得 $p=\dfrac{1}{b-a}$，从而

$$P(\alpha_1<\xi\leq\alpha_2)=\frac{1}{b-a}(\alpha_2-\alpha_1)$$

由此可见，对随机变量均匀分布的情况，只要知道其均匀分布的区间，就可定出 p，这样，这种随机变量的取值规律就完全清楚了．

2　一般情况．我们接下去考虑分布不均匀的情况．我们还是考查概率 $P(\alpha_1<\xi\leq\alpha_2)$．首先把区间 $(\alpha_1,\alpha_2]$ 分细，设细分的小区间长度为 dx．分点为：$\alpha_1=x_1$，$x_2,\cdots,x_n=\alpha_2$．常见的情况是当 dx 非常小时，ξ 的概率分布在这个小区间内的变化幅度也不太大．这样，在每个小区间内，我们都可以近似地把 ξ 的概率分布看成均匀的．即取值于细分的小区间 $(x_i,x_i+dx]$ 内的概率就近似的等于一个常数 p 乘以 dx．又因在每个小区间 p 值不尽相同，所以用 $p(x_i)$ 表示相应于小区间 $(x_i,x_i+dx]$ 上的常数．把上述意思用式子表示出来，即

$$P(x_i<\xi\leq x_i+dx)\doteq p(x_i)dx$$

而

$$P(\alpha_1<\xi\leq\alpha_2)=\sum_{i=1}^{n-1}P(x_i<\xi\leq x_i+dx)$$

$$\doteq\sum_{i=1}^{n-1}p(x_i)dx$$

当 n 越大，dx 越小，$\sum\limits_{i=1}^{n-1}p(x_i)dx$ 就越接近于概率 $P(\alpha_1<\xi\leq\alpha_2)$．特别若令 $n\to\infty$，$dx\to0$ 时，根据积分定义，上述和就是某函数 $p(x)$ 在 $(\alpha_1,\alpha_2]$ 上的积分，即

$$\sum_{i=1}^{n-1}p(x_i)dx\to\int_{\alpha_1}^{\alpha_2}p(x)dx$$

故 $$P(a_1<\xi\leqslant a_2)=\int_{a_1}^{a_2}p(x)dx$$

从上面分析中可以看到，对于连续型随机变量，如果知道了所引出的函数$p(x)$（在均匀情况$p(x)=\begin{cases}\frac{1}{b-a} & a\leqslant x\leqslant b\\ 0 & 其它\end{cases}$），则其全部概率规律性就可以掌握了，随机变量取值于某个区间的概率就等于$p(x)$在这个区间上的积分。对连续型随机变量来说，$p(x)$起着类似离散型随机变量的分布列的作用。因此，为了确切起见，我们就索性用函数$p(x)$来定义连续型随机变量。

定义3.1.1　设ξ是一随机变量，如果存在一个非负可积函数$p(x)$，使得对任意实数$a<b$有

$$P(a<\xi\leqslant b)=\int_a^b p(x)dx$$

则称ξ为连续型随机变量，而$p(x)$称为ξ的分布密度（或称密度函数）。

下面，说明分布密度$p(x)$的意义。分布密度并不表示任何事件的概率，而只表示连续型随机变量概率分布的密集程度。事实上，若我们考虑ξ取值x附近小区间$(x,x+\Delta x)$的概率

$$P(x<\xi\leqslant x+\Delta x)=\int_x^{x+\Delta x}p(x)dx$$

由积分中值定理知道，在$(x,x+\Delta x)$中存在一点x'，使得上式等于$p(x')\Delta x$，而当$\Delta x\to 0$时，它近似等于$p(x)\cdot\Delta x$，即$p(x<\xi\leqslant x+\Delta x)\doteq p(x)\Delta x$。那么如果要问，在$x$附近单位长区间上的概率分布情况怎样呢？这当然应为

$$\frac{P(x<\xi\leqslant x+\Delta x)}{\Delta x}\doteq p(x)$$

取极限 $$\lim_{\Delta x\to 0}\frac{P(x<\xi\leqslant x+\Delta x)}{\Delta x}=p(x)$$

所以$p(x)$实际表示随机变量ξ在x点的概率分布密集程度。

对于分布密度，特别有

$$\int_{-\infty} p(x)dx = P(-\infty<\xi<\infty)=1.$$

对上述内容现给出几何说明，设$p(x)$为ξ的密度函数，则ξ落入区间$(a,b]$的概率$P(a<\xi\leqslant b)=\int_a^b p(x)dx$就是密度曲线下与区间$[a,b]$间的面积（如图3.1.1）。

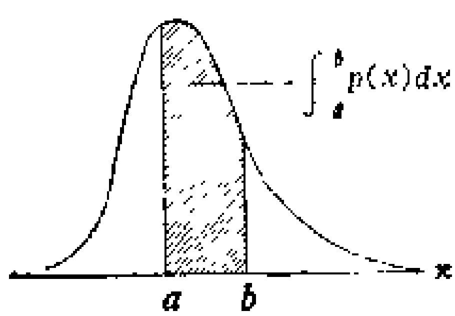


图 （3.1.1）



图 （3.1.2）

例3.1.1　我们每次从手表上看时间时，时针所处的位置ξ显然是在$[0，12]$上均匀分布的随机变量，其密度函数为

$$p(x)=\begin{cases}\dfrac{1}{12} & 0<x\leqslant 12\\ 0 & \text{其他}\end{cases}$$

一般地，在$[a,b]$上均匀分布的随机变量的密度函数为

$$p(x)=\begin{cases}\dfrac{1}{b-a} & a<x<b\\ 0 & \text{其他}\end{cases}$$

例3.1.2　假设连续型随机变量ξ的密度函数$p(x)$型如

$$p(x)=\begin{cases} ax & 0<x<1\\ 0 & \text{其他}\end{cases}$$

求a值，并计算概率$P\left(\xi\leqslant\frac{1}{2}\right)$.

即然$p(x)$为密度函数，故满足$\int_{-\infty}^{\infty}p(x)dx=\int_0^1 axdx=1$，由此推出$a=2$（见图3.1.3）.

$$P\left(\xi\leqslant\frac{1}{2}\right)=\int_{-\infty}^{\frac{1}{2}}p(x)dx=\int_0^{\frac{1}{2}}2xdx=\frac{1}{4}.$$

例3.1.3　设随机变量ξ的密度函数为

$$f(x)=\frac{1}{\sqrt{2\pi}\,\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

其中μ，$\sigma>0$为常数．ξ称为服从正态分布的随机变量．我们可以验证$f(x)$确实是一个密度函数。

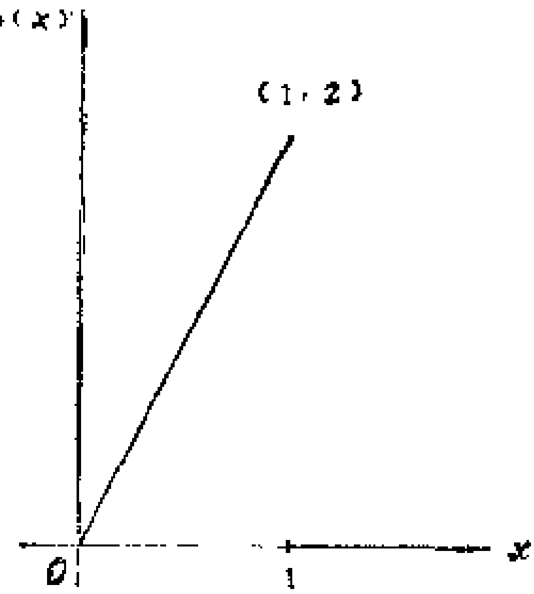


图　(3.1.3)

①　显然$f(x)\geqslant 0$

②　往证$\int_{-\infty}^{\infty}f(x)dx=1$

令　　$t=(x-\mu)/\sigma$，则$I=\int_{-\infty}^{\infty}f(x)dx=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt$

$$I^2=\frac{1}{2\pi}\int_{-\infty}^{\infty}e^{-\frac{t^2}{2}}dt\int_{-\infty}^{\infty}e^{-\frac{s^2}{2}}ds$$

$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}e^{-(s^2+t^2)/2}ds\ dt$$

引进极坐标 $s=r\cos\alpha$　$t=r\sin\alpha$　这时 $dsdt=rdrd\alpha$，随着s，t在$-\infty$到∞之间变化，r在0和∞间变化，α在0到2π间变化．所以

$$I^2=\frac{1}{2\pi}\int_0^{2\pi}\int_0^{\infty}re^{-\frac{r^2}{2}}drd\alpha=\frac{1}{2\pi}\int_0^{2\pi}-e^{-\frac{r^2}{2}}\Big|_0^{\infty}d\alpha$$

$$=\frac{1}{2\pi}\int_0^{2\pi}d\alpha=1$$

正态分布随机变量是一个十分重要的连续型随机变量，在§3.5我们还要详细讨论它．

在实际当中当然还会碰到既不是离散型的又不是连续型的随机变量，不过离散型和连续型随机变量是基础，能够很好地掌握住这两类随机变量，其他类型的随机变量也就不难掌握了．

（二）分布函数

我们知道，分布列刻划了离散型随机变量的概率规律性，分布

密度刻划了连续型随机变量的概率规律性。我们这节介绍的分布函数统一了以上两个概念，它既可以刻划离散型随机变量又可以刻划连续型随机变量的概率规律性。

定义3.1.2　设ξ为随机变量（离散的或连续的），对任意实数x，函数$F(x)=P(\xi\leqslant x)$称为随机变量ξ的分布函数。

定理3.1.1（1）若ξ为一离散型随机变量，其分布列为

$$\begin{pmatrix} x_1, & x_2, & \cdots, & x_n, \cdots \\ p(x_1), & p(x_2), & \cdots, & p(x_n), \cdots \end{pmatrix}$$

则
$$F(x)=\sum_{x_i\leqslant x}p(x_i) \tag{3.1.1}$$

（2）若ξ为连续型随机变量，$p(x)$为其分布密度

则
$$F(x)=\int_{-\infty}^{x}p(u)du \tag{3.1.2}$$

证明　由定义直接推出。

例3.1.4（单点分布与两点分布）。设ξ_1：$\begin{pmatrix} a \\ 1 \end{pmatrix}$，则

$$F_{\xi_1}(x)=\begin{cases} 0 & x<a \\ 1 & x\geqslant a \end{cases}$$

设ξ_2：$\begin{pmatrix} 0, & 1 \\ q, & p \end{pmatrix}$，则其分布函数（见图3.1.4）为

$$F_{\xi_2}(x)=\begin{cases} 0 & x<0 \\ q & 0\leqslant x<1 \\ 1 & 1\leqslant x \end{cases}$$

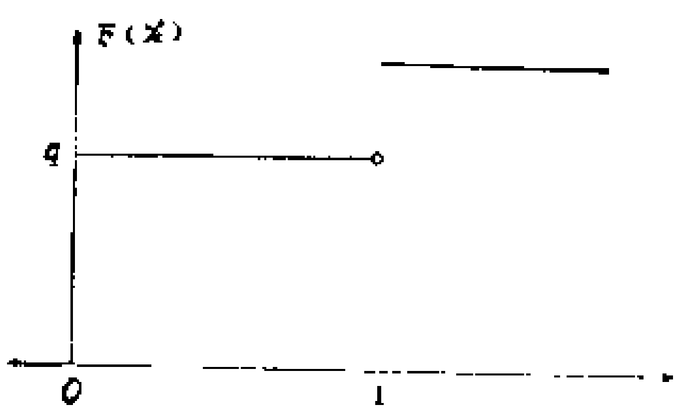


图　(3.1.4)

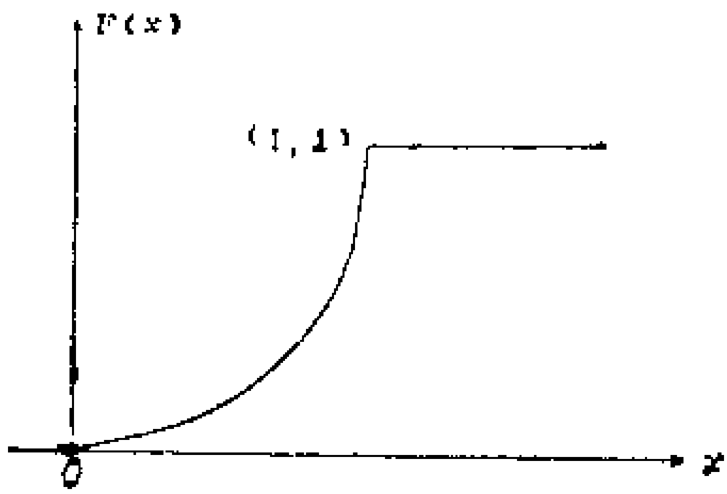
例3.1.5　设ξ为连续型随机变量，其分布密度为

$$p(x)=\begin{cases} 2x & 0<x<1 \\ 0 & \text{其他} \end{cases}$$

则其分布函数为

$$F(x)=\begin{cases}0 & x\leqslant 0\\ \int_0^x 2sds=x^2 & 0<x<1\\ 1 & x\geqslant 1\end{cases}$$

$F(x)$的图形见图（3.1.5）



图（3.1.5）

例3.1.6　设ξ为在$(a,b]$上均匀分布的随机变量，则其分布函数（见图3.1.6）为

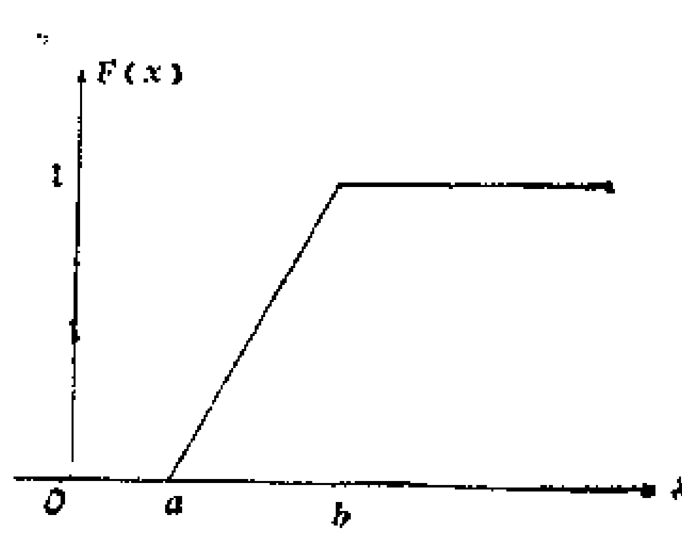
$$F(x)=\begin{cases}0 & x\leqslant a\\ \dfrac{x-a}{b-a} & a<x\leqslant b\\ 1 & x>b\text{时}\end{cases}$$

若ξ是连续型随机变量，由定理3.1.1可知

$$\begin{aligned}P(a<\xi\leqslant b)&=F(b)-F(a)\\&=\int_a^b p(x)dx\end{aligned}$$

而且对任一常数c，$P(\xi=c)=0$。这是因为

$$\begin{aligned}&P(\xi=c)\\&\leqslant P(c-h<\xi\leqslant c)\end{aligned}$$



图（3.1.6）

$$= \int_{c-h}^{c} p(x)dx$$

所以

$$0 \leqslant P(\xi = c) \leqslant \lim_{h \to 0} \int_{c-h}^{c} p(x)dx = 0$$

既然对连续型随机变量来说，它取任何单个值的概率均为0，因此，若ξ是连续型随机变量，则总有$P(a \leqslant \xi \leqslant b) = P(a < \xi \leqslant b) = P(a \leqslant \xi < b) = P(a < \xi < b)$。

定理3.1.2 （1）若ξ是离散型随机变量，则其分布函数$F(x)$为只在ξ的可能取值点$x_1 < x_2 < \cdots < x_n \cdots$上有跳跃的右连续阶梯函数，且

$$p(x_i) \triangleq P(\xi = x_i) = F(x_i) - F(x_{i-1})$$

（2）若ξ是连续型随机变量，则其分布函数连续，且其密度函数

$$p(x) = \frac{d}{dx} F(x)$$

证明（1）假使$x_1 < x_2 < \cdots$，我们有

$$F(x_i) = P\{(\xi = x_i) \cup (\xi = x_{i-1}) \cup \cdots \cup (\xi = x_1)\}$$
$$= p(x_i) + p(x_{i-1}) + \cdots + p(x_1)$$

同样，$F(x_{i-1}) = p(x_{i-1}) + p(x_{i-2}) + \cdots + p(x_1)$

因此 $$F(x_i) - F(x_{i-1}) = p(x_i)$$

（2）显然

例3.1.7 假设连续型随机变量ξ的分布函数为

$$F(x) = \begin{cases} 0 & x \leqslant 0 \\ 1 - e^{-x} & x > 0 \end{cases}$$

则当$x > 0$时$F'(x) = e^{-x}$，因此ξ的分布密度为

$$p(x) = \begin{cases} e^{-x} & x > 0 \\ 0 & \text{其他} \end{cases}$$

定理3.1.3 对任意随机变量ξ，其分布函数$F(x)$有如下性质：

（1）$F(x)$是不减的，即若$x_1<x_2$则有 $F(x_1)\leqslant F(x_2)$

（2）$F(x)$是右连续的，即$F(x)=F(x+0)$

（3）$F(-\infty)\triangleq\lim\limits_{x\to-\infty}F(x)=0$

$$F(\infty)\triangleq\lim_{x\to+\infty}F(x)=1$$

证明（1）因$x_1<x_2$，故事件$A=(\xi\leqslant x_1)\subset B=(\xi\leqslant x_2)$于是有$P(\xi\leqslant x_1)\leqslant P(\xi\leqslant x_2)$

（2）由于$F(x)$的单调性，所以右极限$F(x+0)$存在。为了证明 $F(x+0)=F(x)$，根据单调性，只须证明对某一列$\{x_n\}$，$x_1>x_2>\cdots,x_n\to x$，有$\lim\limits_{n\to\infty}F(x_n)=F(x)$即可。令$A_n=\{x<\xi\leqslant x_n\}$，显然$A_n\supset A_{n+1}$，$\bigcap\limits_{n=1}^{\infty}A_n=\phi$，由连续性定理得

$$\lim_{n\to\infty}F(x_n)-F(x)=\lim_{n\to\infty}[F(x_n)-F(x)]=\lim_{n\to\infty}P(A_n)$$
$$=P(\phi)=0$$

（3）仍然利用单调性，为证$F(-\infty)=0$只证$\lim\limits_{n\to\infty}F(-n)=0$ 即可。考虑$A_n=(\xi\leqslant -n)$，则 $A_n\supset A_{n+1}$，$\bigcap\limits_{n=1}^{\infty}A_n=\phi$，由连续性定理

$$\lim_{n\to\infty}F(-n)=\lim_{n\to\infty}P(A_n)=P(\phi)=0$$

$F(\infty)=1$的证明类似。

定理3.1.4　分布函数$F(x)$的不连续点所成的集是有限集或可列集。

证明　由分布函数的单调不减性，$F(x)$的不连续点集D重合于跳跃点集，因为$0\leqslant F(x)\leqslant 1$，可见跃度不小于$\frac{1}{2^n}$的跳跃点所成的集$D_n$至多只含$2^n$个点，然而，$D=\bigcup\limits_{n=1}^{\infty}D_n$，可见$D$是有限集或可列集。

对离散型随机变量来说，它有几个可能取值，其分布函数就

有几个跳跃点．对连续型随机变量来说，其分布函数是其分布密度积分上限的函数，它是连续的．那么对于既不离散又不连续的随机变量，其分布函数也至多有可列多个跳跃点．

§3.2 连续型随机向量

在第二章我们碰到了离散型随机向量的概念，对于连续型随机变量来说，我们也可以引进连续型随机向量的概念．比如打靶时，每一个弹着点实际上都是一个随机点，它可用弹着点的横坐标ξ和纵坐标η这两个随机变量来表示，它就是一个二维随机变量．本节所研究的随机向量主要是二维连续型随机变量，至于n维情形与二维情形类似．

定义3.2.1　两个随机变量ξ_1，ξ_2的有序整体$\xi=(\xi_1,\xi_2)$称为二维随机变量（或二维随机向量）．对于二维随机向量(ξ_1,ξ_2)，如果存在二元非负可积函数$p(x,y)$，使得对平面R的任一区域Δ，都有

$$P((\xi_1,\xi_2)\in\Delta)=\iint\limits_{\Delta}p(x,y)dxdy \qquad (3.2.1)$$

这里$P((\xi_1,\xi_2)\in\Delta)$表示$(\xi_1,\xi_2)$取值于区域$\Delta$的概率，则称$(\xi_1,\xi_2)$为连续型的．

显然，函数$p(x,y)$满足下列条件：

（1）$p(x,y)\geqslant0$，对一切$(x,y)\subset R$

（2）$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}p(x,y)dxdy=1 \qquad (3.2.2)$$

我们称$p(x,y)$为二维联合分布密度．

特别地，若在(3.2.1)式中，Δ为矩形$a_1\leqslant x\leqslant b_1,a_2\leqslant y\leqslant b_2$时，则

$$P((\xi_1,\xi_2)\in\Delta)=P(a_1\leqslant\xi_1\leqslant b_1,a_2\leqslant\xi_2\leqslant b_2)$$
$$=\int_{a_2}^{b_2}\int_{a_1}^{b_1}p(x,y)dxdy$$

例3.2.1　设(ξ,η)的联合分布密度为

$$p(x,y)=\begin{cases} cy(1-x) & 0\leqslant x\leqslant 1,\ 0\leqslant y\leqslant x \\ 0 & \text{其它} \end{cases}$$

求 c。

解 $1=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}p(x,y)dxdy$

$$=\int_0^1\int_0^x cy(1-x)dydx$$

$$=\int_0^1 c(1-x)\frac{x^2}{2}\ dx=\frac{c}{24}$$

所以 $c=24$

和一维随机变量一样，对二维随机变量可类似定义分布函数。设 (ξ_1,ξ_2) 为二维随机变量，则定义二元函数

$$F(x,y)\triangleq P(\xi_1\leqslant x,\xi_2\leqslant y)$$

为 (ξ_1,ξ_2) 的联合分布函数。特别当 (ξ_1,ξ_2) 为连续型时，设 $p(x,y)$ 为其联合分布密度，则其联合分布函数为

$$F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x}p(u,v)dudv$$

在研究离散型随机变量时我们知道，联合分布能够决定各分量随机变量的分布，连续型随机变量也有同样结果。事实上，$(-\infty<\xi_1<\infty)$，$(-\infty<\xi_2<\infty)$ 均为必然事件，因此，对任意 $a_1<b_1$，

$$P(a_1<\xi_1\leqslant b_1)=P(a_1<\xi_1\leqslant b_1,-\infty<\xi_2<\infty)$$

$$=\int_{-\infty}^{\infty}\int_{a_1}^{b_1}p(x,y,)dxdy$$

$$=\int_{a_1}^{b_1}\left\{\int_{-\infty}^{\infty}p(x,y)dy\right\}dx$$

积分 $\int_{-\infty}^{\infty}p(x,y)dy$ 是参量 x 的函数，以 $p_1(x)$ 表示，则上式可写成

$$P(a_1<\xi_1\leqslant b_1)=\int_{a_1}^{b_1}p_1(x)dx \tag{3.2.3}$$

同样对任意 $a_2<b_2$，有

$$P(a_2<\xi_2\leqslant b_2)=\int_{a_2}^{b_2}p_2(y)dy \tag{3.2.4}$$

其中$p_2(y)=\int_{-\infty}^{\infty}p(x,y)dx$。我们称 $p_1(x)$和$p_2(y)$为(ξ_1,ξ_2)的边缘分布密度。(3.2.3)和(3.2.4)式说明边缘分布密度$p_1(x)$和 $p_2(y)$恰好分别是ξ_1与ξ_2的分布密度。因此，如果知道了联合分布密度，就可以推出各分量随机变量的分布密度。

例3.2.2　设(ξ,η)为例3.2.1中给出的随机向量，即其联合分布密度为

$$p(x,y)=24y(1-x)\qquad 0\leqslant x\leqslant 1,\ 0\leqslant y\leqslant x$$

分布密度的定义范围见图(3.2.1)。

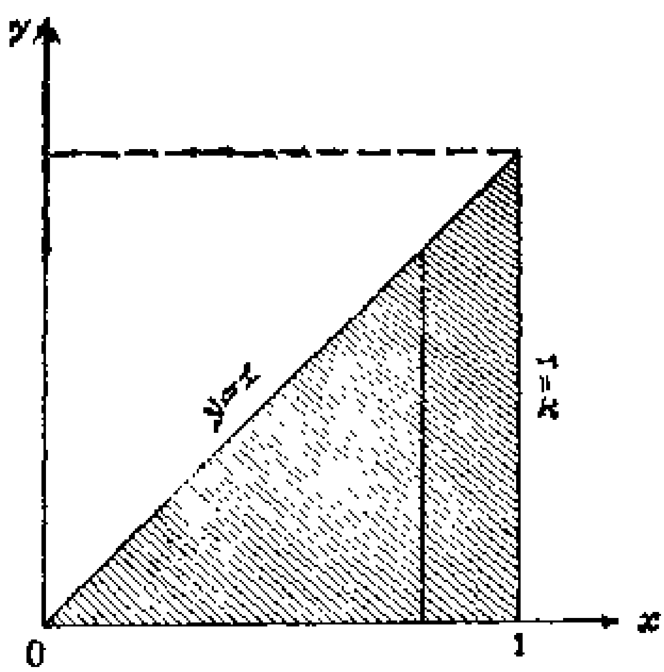


图　(3.2.1)

则其边缘分布密度为

$$p_1(x)=\int_{-\infty}^{\infty}p(x,y)dy=24(1-x)\int_0^x ydy$$

$$=12x^2(1-x)\qquad (0\leqslant x\leqslant 1)$$

$$p_2(y)=\int_{-\infty}^{\infty}p(x,y)dx=24y\int_y^1(1-x)dx$$

$$=12y(1-y)^2\qquad (0\leqslant y\leqslant 1)$$

例3.2.3　火箭发动机性能的两个特性是推力ξ和混合比　η。假设(ξ,η)为二维连续型随机变量，其联合分布密度为

$$p(x,y)=\begin{cases}2(x+y-2xy) & 0\leqslant x\leqslant 1,\ 0\leqslant y\leqslant 1\\ 0 & \text{其它}\end{cases}$$

则 ξ 的边缘分布密度为

$$p_1(x)=\int_0^1 2(x+y-2xy)dy=2\left(xy+\frac{y^2}{2}-xy^2\right)\Big|_0^1$$

$$=1 \qquad 0\leqslant x\leqslant 1$$

这说明 ξ 是在$[0,1]$上均匀分布的随机变量。

η 的边缘分布密度为

$$p_2(y)=\int_0^1 2(x+y-2xy)dx=1,\ 0\leqslant y\leqslant 1.$$

因此 η 也是在$[0,1]$上均匀分布的随机变量。

现在来考虑相反的问题，是否知道了两个边缘分布就能确定其联合分布呢？与离散型情况一样在一般情况下是做不到的，而只有在两个随机变量是独立的情况下，回答才是肯定的。

定义3.2.2 (ξ_1,ξ_2)为连续型随机向量，若

$$p(x,y)=p_1(x)\cdot p_2(y)$$

其中$p(x,y)$、$p_1(x)$、$p_2(y)$分别是(ξ_1,ξ_2)、ξ_1与ξ_2的密度函数，则称随机变量ξ_1与ξ_2是相互独立的。

例3.2.4 设 ξ 和 η 是两个电子设备的寿命，它们的联合分布密度为

$$p(x,y)=e^{-(x+y)} \qquad x\geqslant 0,\ y\geqslant 0$$

由于 $p(x,y)=e^{-x}\cdot e^{-y}$，而$e^{-x}$和 e^{-y}又都是密度函数，所以可断言 ξ 与 η 是相互独立的。注意，并不是当联合分布密度可以写成 x 的函数与 y 的函数乘积形式时，ξ 与 η 就是独立的，比如在例3.2.2里，其中联合分布密度尽管可写成 x、y 的函数乘积，但 ξ 与 η 并不独立，因为分解的因子不是密度函数。

定理3.2.1 ξ 与 η 相互独立的充分必要条件是对任意 $a<b$，$c<d$，事件$\{a<\xi<b\}$与$\{c<\eta<d\}$相互独立。

证明 必要性，设 $p(x,y)=p_1(x)\cdot p_2(y)$是$(\xi,\eta)$的联合密度，则

$$P(\{a<\xi<b\}\cdot\{c<\eta<d\})$$
$$=\iint\limits_{\substack{a<x<b\\c<y<d}} p(x,y)dxdy$$
$$=\int_a^b p_1(x)dx\cdot\int_c^d p_2(y)dy$$
$$=P(a<\xi<b)\cdot P(c<\eta<d)$$

充分性 设$\Delta=\{(x,y):a<x<b,c<y<d\}$，则

$$P((\xi,\eta)\in\Delta)=\iint\limits_{\Delta}p(x,y)dxdy$$

另一方面

$$\begin{aligned}P((\xi,\eta)\in\Delta)&=P(a<\xi<b,c<\eta<d)\\&=P(a<\xi<b)\cdot P(c<\eta<d)\\&=\int_a^b p_1(x)dx\cdot\int_c^d p_2(y)dy\\&=\iint\limits_{\Delta}p_1(x)\cdot p_2(y)dxdy\end{aligned}$$

因此推得 $p(x,y)=p_1(x)\cdot p_2(y)$.

下面仿照二维情况来对n维随机向量进行扼要的讨论.

定义3.2.3 有序的n个随机变量$(\xi_1,\xi_2,\cdots,\xi_n)$的总体称为$n$维随机向量，如果对于$n$维随机向量$(\xi_1,\xi_2,\cdots,\xi_n)$来说，存在$n$元非负可积函数$p(x_1,x_2,\cdots,x_n)$，使得对$n$维空间的任一区域$\Delta$，都有

$$P((\xi_1,\xi_2,\cdots,\xi_n)\in\Delta)=\int\cdots\int\limits_{\Delta}p(x_1,x_2,\cdots,x_n)dx_1dx_2\cdots dx_n$$

则称$(\xi_1,\xi_2,\cdots,\xi_n)$是连续型的，函数$p(x_1,x_2,\cdots,x_n)$称为$(\xi_1,\xi_2,\cdots,\xi_n)$的联合分布密度.

特别地，当Δ为n维长方体$D=\{(x_1,\cdots,x_n):a_1<x_1<b_1,\cdots,a_n<x_n<b_n\}$时，有

$$P((\xi_1,\xi_2\cdots,\xi_n)\in D)=\int_{a_n}^{b_n}\cdots\int_{a_1}^{b_1}p(x_1,x_2,\cdots,x_n)dx_1\cdots dx_n$$

同二维情形一样，知道联合分布密度，其边缘分布也可以求出。比如ξ_1的边缘分布密度为

$$p_1(x_1)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}p(x_1,x_2,\cdots,x_n)dx_2\cdots dx_n$$

(ξ_1,ξ_2)的边缘分布密度为

$$p(x_1,x_2)=\int_{-\infty}^{\infty}\cdots\int_{-\infty}^{\infty}p(x_1,x_2\cdots,x_n)dx_3\cdots dx_n$$

定义3.2.4　设$\xi_1,\xi_2,\cdots,\xi_n$的联合分布密度为$p(x_1,x_2,\cdots,x_n)$，其边缘分布密度分别是$p_1(x_1),\cdots,p_n(x_n)$，如果$p(x_1,x_2,\cdots,x_n)=p_1(x_1)p_2(x_2)\cdots p_n(x_n)$，则称$\xi_1,\xi_2,\cdots,\xi_n$是相互独立的。

容易验证，相互独立的充要条件是事件$(a_1<\xi_1<b_1),(a_2<\xi_2<b_2),\cdots,(a_n<\xi_n<b_n)$相互独立，其中$a_1<b_1,\cdots,a_n<b_n$为任意$n$对数。

§3.3　连续型随机变量的函数

（一）一元函数

在研究随机变量时，常常会碰到以随机变量为自变量的函数，比如ξ为一随机变量，$H(x)$为一函数，则$H(\xi)$就是以ξ为自变量的函数。在这类随机变量函数中，最重要的也是最常碰到的是这样一种特殊情形，即ξ是连续型随机变量，函数$H(x)$是一连续函数，这时随机变量$H(\xi)$也是一个连续型随机变量。问题是当给定了$H(x)$的解析表达式后，如何根据ξ的分布，求出$\eta=H(\xi)$的分布。计算η的密度一般可分下面三步。

（1）设法根据ξ的分布计算出η的分布函数$G(y)$。

（2）对$G(y)$求导数得到$g(y)$。

（3）在η的值域里，确定出使$g(y)>0$的y值。

例3.3.1　设ξ的分布密度为

$$f(x)=\begin{cases}2x & 0<x<1\\ 0 & 其它\end{cases}$$

$H(x)=3x+1$（见图3.3.1），求 $\eta=H(\xi)=3\xi+1$ 的分布密度。

解 $G(y)=P(\eta\leqslant y)=P(3\xi+1\leqslant y)$

$$=P\left(\xi\leqslant\frac{y-1}{3}\right)=\int_0^{\frac{y-1}{3}}2x\,dx$$

$$=[(y-1)/3]^2$$

于是

$$g(y)=G'(y)=\frac{2}{9}(y-1)$$

因为 当 $0<x<1$ 时，$f(x)>0$，所以当 $0<\frac{y-1}{3}<1$，即 $1<y<4$ 时 $g(y)>0$

例3.3.2 设 ξ 同上例，$H(x)=e^{-x}$（图3.3.2），求 $\eta=H(\xi)=e^{-\xi}$ 的分布密度。

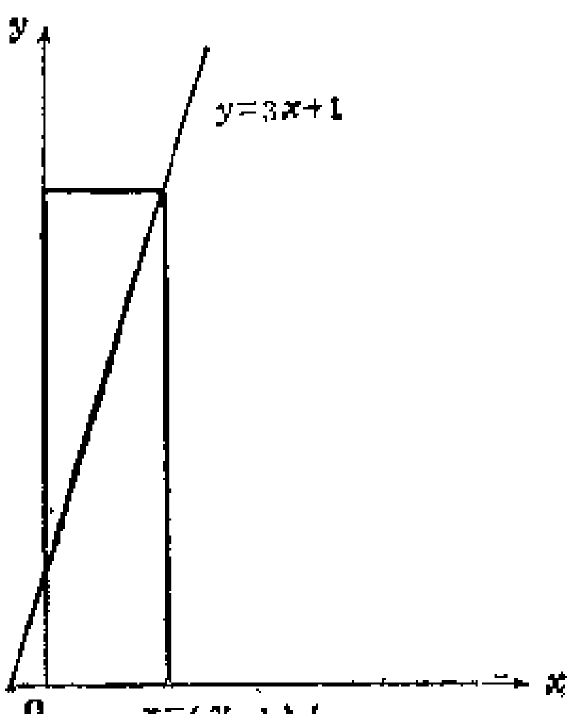


图 (3.3.1)

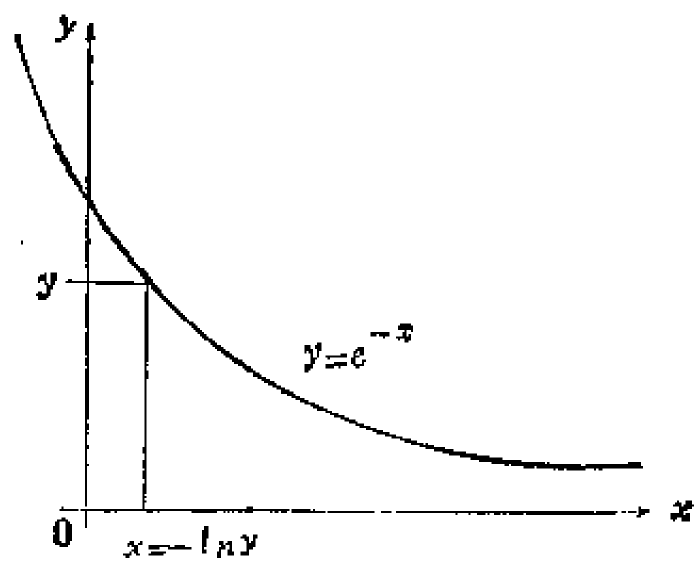


图 (3.3.2)

解 $G(y)=P(\eta\leqslant y)=P(e^{-\xi}\leqslant y)$

$$=P(\xi\geqslant-\ln y)=\int_{-\ln y}^{\infty}p(x)\,dx=\int_{-\ln y}^{1}2x\,dx$$

$$=1-(-\ln y)^2$$

$$g(y)=G'(y)=-2\frac{\ln y}{y}$$

因为当$0<x<1$时$p(x)>0$，所以当$0<-\ln y<1$，即$0>\ln y>-1$或$\frac{1}{e}<y<1$时，$g(y)>0$。

对于$H(x)$为单调函数的情形，我们有下面定理。

定理3.3.1 设ξ是连续型随机变量，其分布密度为$p(x)$，（$p(x)>0$，当$a<x<b$时）。又设$y=H(x)$为x的严格单调（增或减）可微函数，则随机变量的函数$\eta=H(\xi)$的分布密度$g(y)$可如下求出

$$g(y)=p(x)\left|\frac{dx}{dy}\right| \tag{3.3.1}$$

其中$x=H^{-1}(y)$。当H是递增时，则$g(y)$在区间$H(a)<y<H(b)$不为零，当H是递减时，$g(y)$在区间$H(b)<y<H(a)$不为零。

证明 （1）假设H是严格增函数，则

$$G(y)=P(\eta<y)=P(H(\xi)\leqslant y)$$

$$=P(\xi<H^{-1}(y))=F(H^{-1}(y))$$

$$\frac{dG(y)}{dy}=\frac{dG(y)}{dx}\cdot\frac{dx}{dy} \qquad 其中x=H^{-1}(y)$$

于是$G'(y)=\frac{dF(x)}{dx}\frac{dx}{dy}=f(x)\frac{dx}{dy}$。

（2）假设H是减函数，

则 $$G(y)=P(\eta\leqslant y)=P(H(\xi)\leqslant y)=P(\xi\geqslant H^{-1}(y))$$

$$=1-F(H^{-1}(y))$$

$$\frac{dG(y)}{dy}=\frac{dG(y)}{dx}\frac{dx}{dy}=\frac{d}{dx}[1-F(x)]\frac{dx}{dy}$$

$$=-f(x)\frac{dx}{dy}$$

例3.3.3 现在利用定理3.3.1对上面两个例子进行计算：

（1）在例3.3.1中，$f(x)=2x$，当$0<x<1$，因函数$y=3x+1$

单调增，并且 $x=(y-1)/3$，$dx/dy=\frac{1}{3}$，所以，由定理3.3.1

$$g(y)=2[(y-1)/3]\frac{1}{3}=\frac{2}{9}(y-1),\quad 1<y<4.$$

（2）在例3.3.2中，$f(x)=2x$，$0<x<1$，且$y=e^{-x}$单调减，$x=-\ln y$，$dx/dy=-1/y$，因此 $g(y)=-2(\ln y)/y$，$1/e<y<1$，上面的结果与例3.3.1与例3.3.2中的计算结果完全相同。

例3.3.4　设 ξ 为非负连续型随机变量，其分布密度为 $p(x)$，又给定$H(x)=x^a$，其中$a>0$，求随机变量$\eta=H(\xi)=\xi^a$的分布密度。

设 η 的分布密度为$g(y)$。因为$H(x)$在$[0,\infty)$里是单调可微的，由定理3.3.1我们有

$$g(y)=p(y^{\frac{1}{a}})\left|\frac{dy^{\frac{1}{a}}}{dy}\right|$$

$$=\begin{cases}\dfrac{1}{a}y^{\frac{1}{a}-1}p(y^{\frac{1}{a}}) & y>0\\ 0 & y\leqslant 0\end{cases}$$

例3.3.5　设 ξ 在$[0,\pi]$中均匀分布，$H(x)=\sin x$，试求 $\eta=\sin\xi$ 的分布函数$G(y)$

因 $H(x)$在$[0,\pi]$区间上不是单调函数，故不能用定理3.3.1计算。但由于 $\xi\in[0,\pi]$，故 η 只能取$[0,1]$中的值，对任意$y\in[0,1]$，由图(3.3.3)知

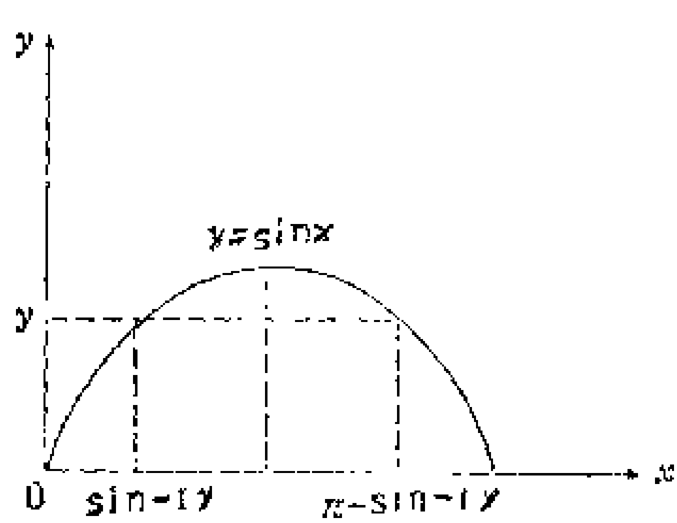


图　(3.3.3)

$$
\begin{aligned}
G(y) &= P(\sin\xi \leqslant y) \\
&= P(0 \leqslant \xi \leqslant \sin^{-1} y) + P(\pi - \sin^{-1} y \leqslant \xi \leqslant \pi) \\
&= \frac{1}{\pi}[\sin^{-1} y + \pi - (\pi - \sin^{-1} y)] \\
&= \frac{2\sin^{-1} y}{\pi}
\end{aligned}
$$

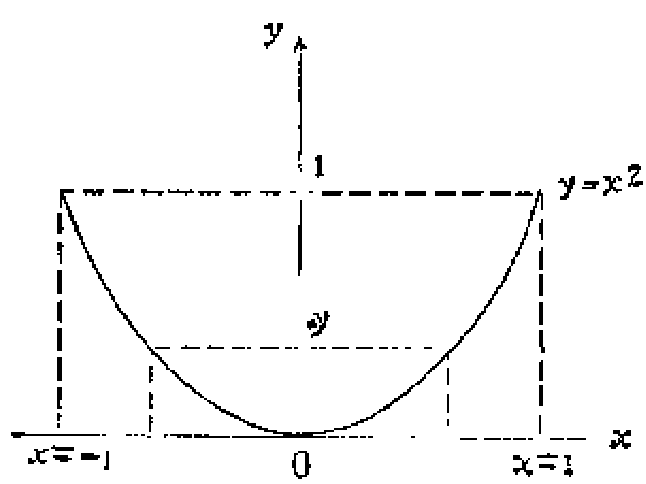
因此

$$
G(y) = \begin{cases} 0 & y \leqslant 0 \\ \dfrac{2\sin^{-1} y}{\pi} & 0 \leqslant y \leqslant 1 \\ 1 & y > 1 \end{cases}
$$

其密度函数是

$$
g(y) = \frac{2}{\pi\sqrt{1-y^2}}, \quad 0 \leqslant y \leqslant 1.
$$

例3.3.6　设 ξ 在$[-1, 1]$中均匀分布，$H(x) = x^2$，求$\eta = \xi^2$的分布。由函数$H(x) = x^2$的图象（图3.3.4）可知，若$\xi \in [-1, 1]$，η 取值于$[0,1]$，所以当$0 < y < 1$时，



图（3.3.4）

$$
\begin{aligned}
G(y) &= P(\eta \leqslant y) = P(\xi^2 \leqslant y) \\
&= P(-\sqrt{y} \leqslant \xi \leqslant \sqrt{y}) \\
&= F(\sqrt{y}) - F(-\sqrt{y})
\end{aligned}
$$

其中 $F(\cdot)$ 为 ξ 的分布函数，因此

$$g(y)=G'(y)=\frac{f(\sqrt{y})}{2\sqrt{y}}+\frac{f(-\sqrt{y})}{2\sqrt{y}}$$

$$=\frac{1}{2\sqrt{y}}\left[f(\sqrt{y})+f(-\sqrt{y})\right]$$

因此 $$g(y)=\frac{1}{2\sqrt{y}}\left(\frac{1}{2}+\frac{1}{2}\right)=\frac{1}{2\sqrt{y}},\qquad 0<y<1.$$

请注意，上面两个例子中的函数$H(x)$并不是单调的，而是“分段单调的”．我们可以把定理3.3.1扩充到这样分段单调函数上．

定理3.3.2　设ξ是连续型随机变量，其分布密度为$p(x)$．令$y=H(x)$是分段单调函数，即在不相重叠的区间$I_1,I_2,\cdots$上逐段严格单调，其反函数$w_1(y),w_2(y),\cdots$以及其导函数$w_1'(y),w_2'(y),\cdots$均为连续函数，则$\eta=H(\xi)$的密度函数

$$g(y)=\sum_i p(w_i(y))\,|w_i'(y)| \tag{3.3.2}$$

证明　给定实数a，以$E_i(a)$记I_i中使$H(x)\leqslant a$成立的x的集

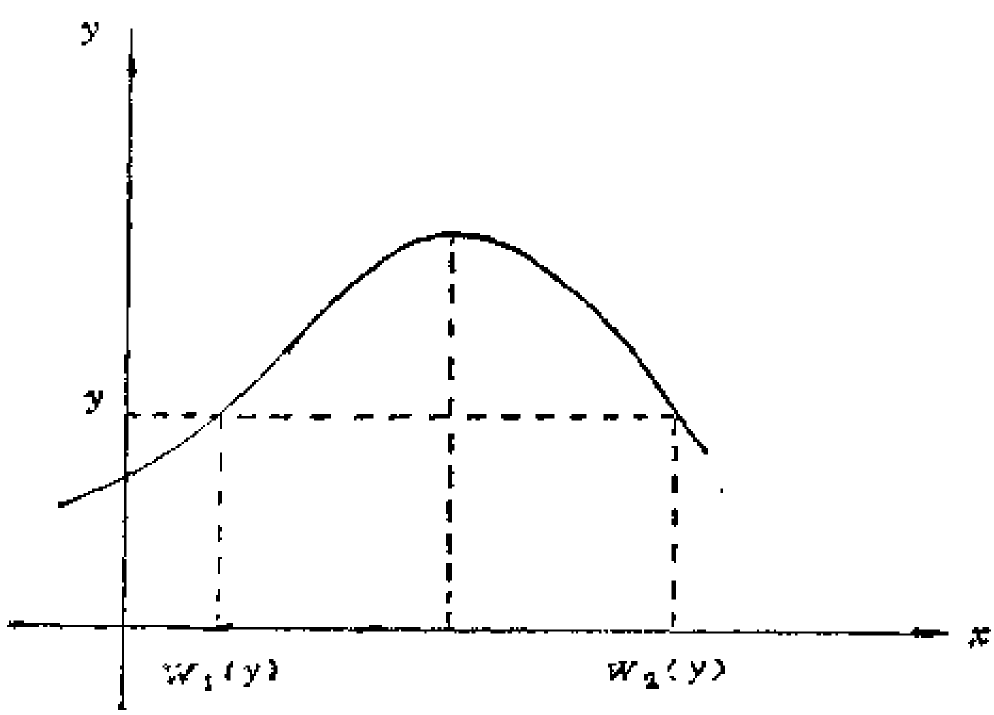


图 (3.3.5)

合，显然诸$E_i(a)$不相交，而

$$P(\eta\leqslant a)=P(H(\xi)\leqslant a)=P(\xi\in\bigcup_i E_i(a))$$

$$= \sum_i \int_{I_i(y)} p(x)dx$$

$$= \sum_i \int_{-\infty}^{y} p(w_i(y))|w_i'(y)|dy$$

$$= \int_{-\infty}^{y} \sum_i p(w_i(y))|w_i'(y)|dy$$

因此η的密度函数为$g(y)=\sum_i p(w_i(y)|w_i'(y)|$。

（二）二维随机变量的函数

下面我们要讨论的问题是假如(ξ,η)是二维连续型随机变量，$H(x,y)$是一个二元函数，如果已知(ξ,η)的分布，如何求新随机变量$\zeta=H(\xi,\eta)$的分布。问题的更一般的提法是，若(ξ,η)是连续型二维随机向量，$H_1(x,y)$，$H_2(x,y)$是两个二元函数，而$\zeta=H_1(\xi,\eta)$，$\omega=H_2(\xi,\eta)$，那么如何根据(ξ,η)的联合分布求出(ζ,ω)的联合分布？

定理3.3.3　设(ξ,η)是二维连续型随机向量，其联合分布密度为$p(x,y)$。令$z=H_1(x,y)$，$w=H_2(x,y)$是平面到自身的一对一映射，其逆映射为$x=G_1(z,w)$，$y=G_2(z,w)$。又假设上述映射和逆映射都是连续的，偏导数$\partial x/\partial z,\partial x/\partial w,\partial y/\partial z,\partial y/\partial w$都存在且连续，则二维随机向量$(\zeta,\omega)$，其中$\zeta=H_1(\xi,\eta)$，$\omega=H_2(\xi,\eta)$的联合分布密度为

$$f(z,w)=P(G_1(z,w),G_2(z,w))|J(z,w)| \qquad (3.3.3)$$

其中$J(z,w)$是逆变换的雅可比(Jacobian)行列式

$$J(z,w)=\frac{\partial(x,y)}{\partial(z,w)}=\begin{vmatrix} \dfrac{\partial x}{\partial z} & \dfrac{\partial x}{\partial w} \\ \dfrac{\partial y}{\partial z} & \dfrac{\partial y}{\partial w} \end{vmatrix}$$

证明　设D是$z-w$平面上的区域，D^*为由$\begin{cases} x=G_1(z,w) \\ y=G_2(z,w) \end{cases}$，其中$(z,w)\subset D$所规定的$x-y$平面的区域。

则有

$$P((\zeta,\omega)\in D)=P((\xi,\eta)\in D^*)$$

$$=\iint_{D^*} p(x,y)dxdy$$

作变量替换$\begin{cases}z=H_1(x,y)\\w=H_2(x,y)\end{cases}$ $\left(\text{逆变换为}\begin{array}{l}x=G_1(z,w)\\y=G_2(z,w)\end{array}\right)$

$$\iint_{D^*} p(x,y)dxdy=\iint_{D} p(G_1(z,w),G_2(z,w))|J|dzdw$$

因此有 $p((\zeta,\omega)\in D)=\iint_D p(G_1(z,w),G_2(z,w)|J|dzdw$

这表明(ζ,w)的联合分布密度为

$$p(G_1(z,w),\ G_2(z,w))|J|.$$

例3.3.7 设ξ，η相互独立同均匀分布，分布密度为

$$p(x)=\begin{cases}1 & \text{当}0<x<1\\0 & \text{其他}\end{cases}$$

令$\zeta=\xi+\eta$，$\omega=\xi-\eta$，求(ζ,ω)的联合分布。

解 我们有变换$\begin{cases}z=x+y\\w=x-y\end{cases}$，逆变换为$\begin{cases}x=\dfrac{z+w}{2}\\y=\dfrac{z-w}{2}\end{cases}$

$$J(z,w)=\begin{vmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & -\frac{1}{2}\end{vmatrix}=-\frac{1}{2}$$

所以

(ξ,ω)的联合分布密度为

$$f(z,w)=\frac{1}{2}p\left(\frac{z+w}{2}\right)p\left(\frac{z-w}{2}\right)$$

$$=\begin{cases}\frac{1}{2} & \text{当}0<\frac{z+w}{2}<1,\ 0<\frac{z-w}{2}<1\\0 & \text{其它}\end{cases}$$

例3.3.8　设ξ，η是两个连续型随机变量，其联合分布密度为

$$p(x,y)=\begin{cases}4xy & 0<x<1,\ 0<y<1\\ 0 & \text{其它}\end{cases}$$

求　$\zeta=\xi^2$，$\omega=\xi\eta$的联合分布密度。

变换　$z=x^2$，$w=xy$的逆变换为$x=\sqrt{z}$，$y=\frac{w}{\sqrt{z}}$，由此我们得到

$$J(z,w)=\begin{vmatrix}\frac{1}{2\sqrt{z}} & 0\\ -\frac{w}{2z^{\frac{3}{2}}} & \frac{1}{\sqrt{z}}\end{vmatrix}=\frac{1}{2z}$$

(ζ,ω)的联合分布密度为

$$f(z,w)=P\left(\sqrt{z},\ \frac{w}{\sqrt{z}}\right)\left|\frac{1}{2z}\right|$$

$$=\begin{cases}4\sqrt{z}\,\frac{w}{\sqrt{z}}\cdot\frac{1}{2z} & \text{当}0<\sqrt{z}<1,\ 0<\frac{w}{\sqrt{z}}<1\\ 0 & \text{其它}\end{cases}$$

$$=\begin{cases}\frac{2w}{z} & \text{当}0<\sqrt{z}<1,\ 0<w<\sqrt{z}\\ 0 & \text{其它}\end{cases}$$

利用定理3.3.3，我们可以根据两个随机变量的联合分布密度，求出它们的和、差、积、商的分布。

定理3.3.4　设$p(x,y)$是连续型随机向量(ξ,η)的联合分布密度，令

$$Z=\xi+\eta$$
$$U=\xi-\eta$$
$$V=\xi\cdot\eta$$

若$P(\eta=0)=0$时，令

$$W=\xi/\eta$$

则Z，U，V，W的分布密度分别可如下求出

$$p_Z(z) = \int_{-\infty}^{\infty} p(x, z-x)dx \tag{3.3.4}$$

$$p_U(u) = \int_{-\infty}^{\infty} p(u+x, x)dx \tag{3.3.5}$$

$$p_V(v) = \int_{-\infty}^{\infty} p\left(x, \frac{v}{x}\right)\frac{1}{|x|}dx \tag{3.3.6}$$

$$p_W(w) = \int_{-\infty}^{\infty} p(xw, x)|x|dx \tag{3.3.7}$$

证明　为证(3.3.4)式，令$z = x + y, w = x$，于是$x = w, y = z - w$，

$$J(z,w) = \begin{vmatrix} 0 & 1 \\ 1 & -1 \end{vmatrix} = -1$$

由定理3.3.3，随机向量(ξ, Z)的联合分布密度为

$$f(w,z) = p(w, z-w)$$

其中$Z = \xi + \eta$，求Z的边缘分布

$$p_Z(z) = \int_{-\infty}^{\infty} f(w,z)dw = \int_{-\infty}^{\infty} p(w, z-w)dw.$$

下面证（3.3.6）式。令　$v = xy$，$w = x$，于是　$x = w, y = \dfrac{v}{w}$，

$$J(v,w) = \begin{vmatrix} 0 & 1 \\ \dfrac{1}{w} & -\dfrac{v}{w^2} \end{vmatrix} = -\frac{1}{w}$$，则(ξ, V)的联合分布密度为

$$h(w,v) = p\left(w, \frac{v}{w}\right)\left|\frac{1}{w}\right|$$

其中$V = \xi\cdot\eta$，因此，V的分布密度

$$p_V(v) = \int_{-\infty}^{\infty} p\left(w, \frac{v}{w}\right)\left|\frac{1}{w}\right|dw$$

(3.3.5)和(3.3.7)式的证明与（3.3.4）、（3.3.6）式的证明类似，留给读者做为练习。

推论　若ξ，η相互独立，分布密度分别为$p_1(x)$和$p_2(y)$，则和、差、积、商的分布密度分别为

$$p_Z(z) = \int_{-\infty}^{\infty} p_1(x)p_1(z-x)dx \tag{3.3.8}$$

$$p_U(u) = \int_{-\infty}^{\infty} p_1(u+x)p_2(x)dx \tag{3.3.9}$$

$$p_V(v) = \int_{-\infty}^{\infty} p_1(x) p_2\left(\frac{v}{x}\right) \frac{1}{|x|} dx \qquad (3.3.10)$$

$$p_W(w) = \int_{-\infty}^{\infty} p_1(xw) p_2(x) |x| dx. \qquad (3.3.11)$$

例3.3.9　ξ、η相互独立同在$[0,1]$上均匀分布，求$\xi+\eta$的分布密度。

解　ξ与η的分布密度均为

$$p(x) = \begin{cases} 1 & 0 \leqslant x \leqslant 1 \\ 0 & \text{其它} \end{cases}$$

利用定理3.3.4，$Z=\xi+\eta$的分布密度

$$p_Z(z) = \int_{-\infty}^{\infty} p(x) p(z-x) dx$$

因为被积函数只有在$0 \leqslant x \leqslant 1$，$0 \leqslant z-x \leqslant 1$时非零，而$\begin{cases} 0 \leqslant x \leqslant 1 \\ 0 \leqslant z-x \leqslant 1 \end{cases}$所确定的$x-z$平面的区域为如图（3.3.6）中阴影部分，故

$$p_Z(z) = \begin{cases} \int_0^z 1dx & 0 \leqslant z \leqslant 1 \\ \int_{z-1}^1 1dx & 1 < z \leqslant 2 \\ 0 & \text{其它} \end{cases}$$

$$= \begin{cases} z & 0 \leqslant z \leqslant 1 \\ 2-z & 1 < z \leqslant 2 \\ 0 & \text{其它} \end{cases}$$

$p_Z(z)$的图形见图（3.3.7）。

例3.3.10　设电路里电流I和电阻R都是随机变量，它们相互独立，其分布密度分别为

$$I: g(i) = \begin{cases} 2i & 0 \leqslant i \leqslant 1 \\ 0 & \text{其它} \end{cases}$$

$$R: h(r) = \begin{cases} r^2/9 & 0 \leqslant r \leqslant 3 \\ 0 & \text{其它} \end{cases}$$

求$E=IR$的分布密度。

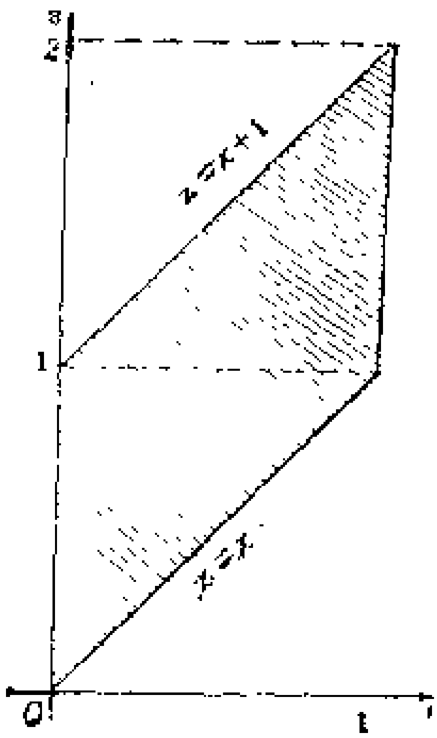


图 (3.3.6)

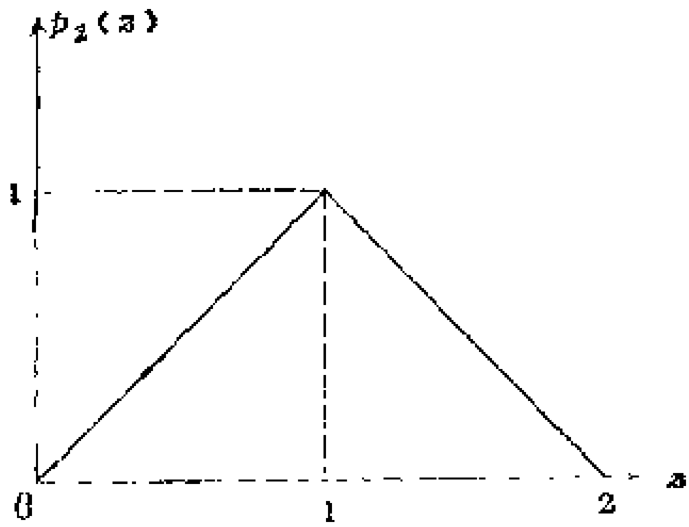


图 (3.3.7)

解　由定理3.3.4，E的分布密度为

$$p(e)=\int_{-\infty}^{\infty}g(i)h\left(\frac{e}{i}\right)\left|\frac{1}{i}\right|di$$

当$0\leqslant i\leqslant 1$，$0\leqslant e/i\leqslant 3$时（图(3.3.8)中的阴影部分），上式的被积函数不为零。因此，对满足$0\leqslant e\leqslant 3$式的e，有

$$p(e)=\int_{e/3}^{1}2i\frac{e^2}{9i^2}\cdot\frac{1}{i}di$$

$$=-\frac{2}{9}e^2\frac{1}{i}\Big|_{e/3}^{1}$$

$$=\frac{2}{9}e(3-e)$$

函数$p(e)$的形状见图（3.3.9）。

有些随机变量的二元函数的分布直接利用其分布函数的定义去计算也很方便。

例3.3.11　设ξ,η相互独立且有相同的分布密度$\frac{1}{\sqrt{2\pi}}exp\left(-\frac{x^2}{2}\right)$ $(-\infty<x<\infty)$求$\zeta=\sqrt{\xi^2+\eta^2}$的分布密度。

解　ζ的分布函数

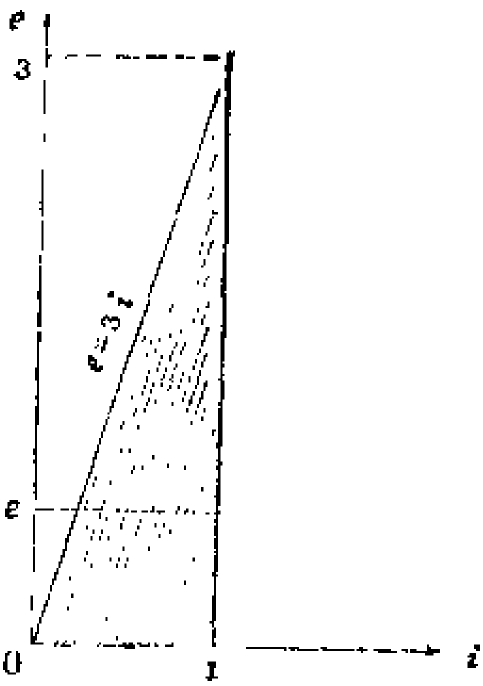


图 (3.3.8)

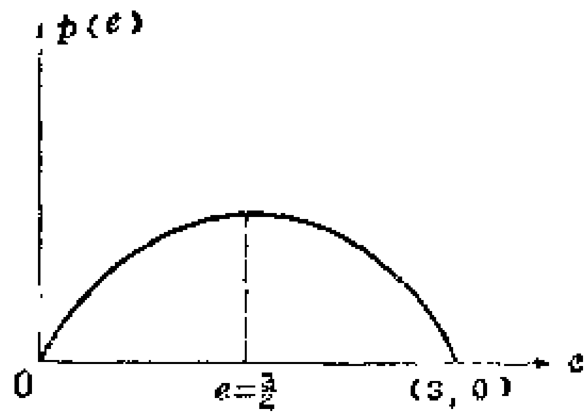


图 (3.3.9)

$$F_\zeta(z)=P(\zeta\leqslant z)=P(\sqrt{\xi^2+\eta^2}\leqslant z)$$

$$=\iint\limits_{\sqrt{x^2+y^2}\leqslant z}\frac{1}{2\pi}e^{-\frac{1}{2}(x^2+y^2)}dxdy$$

作极坐标变换$x=r\cos\theta$，$y=r\sin\theta$（$r\geqslant 0$，$0\leqslant\theta<2\pi$）。当$z>0$时，

$$F_\zeta(z)=\int_0^{2\pi}d\theta\int_0^z\frac{r}{2\pi}e^{-\frac{1}{2}r^2}dr=\int_0^z re^{-\frac{1}{2}r^2}dr$$

当$z\leqslant 0$时，$F_\zeta(z)=0$。因此，ζ的分布密度

$$p_\zeta(z)=\begin{cases} ze^{-\frac{z^2}{2}} & z>0\\ 0 & \text{其它}\end{cases}$$

这就是所谓的瑞利(Rayleigh)分布。

下面我们讨论作为n个随机变量的函数的所谓极值分布问题。

定理3.3.5 假设随机变量$\xi_1,\xi_2,\cdots,\xi_n$相互独立同分布（设其共同的分布函数为F(x)，分布密度为$p(x)$）。令$\xi_1^*=\min\limits_{1\leqslant i\leqslant n}\{\xi_i\}$，$\xi_n^*=\max\limits_{1\leqslant i\leqslant n}\{\xi_i\}$，则$\xi_1^*$和$\xi_n^*$的分布函数分别为

$$G_1(x)=1-[1-F(x)]^n$$

$$G_n(x)=[F(x)]^n$$

证明 $G_1(x)=P(\xi_1^*\leqslant x)$

$$=P(\text{在}\ \xi_1,\xi_2,\cdots,\xi_n\ \text{中至少有一个小于等于}\ x)$$

$$=P\left(\bigcup_{i=1}^{n}(\xi_i\leqslant x)\right)=1-P\left(\bigcap_{i=1}^{n}(\xi_i>x)\right)$$

$$=1-[1-F(x)]^n$$

$$G_n(x)=P(\xi_n^*\leqslant x)=P\left(\bigcap_{i=1}^{n}(\xi_i\leqslant x)\right)$$

$$=[F(x)]^n$$

设ξ_1^*，ξ_n^*的分布密度分别记为$g_1(x)$和$g_n(x)$，则

$$g_1(x)=G_1'(x)=n[1-F(x)]^{n-1}p(x)$$

$$g_n(x)=G_n'(x)=n[F(x)]^{n-1}p(x)$$

§3.4 连续型随机变量的数字特征

对于连续型随机变量也有刻划它的平均状态和取值离散程度的量，这就是连续型随机变量的均值和方差。我们已经掌握了离散型随机变量的均值和方差，下面我们就联系离散情况来讨论连续型随机变量的均值和方差应如何定义。

（一）均　　值

在离散情况，均值定义为 $E\xi=\sum_{i=1}^{n}x_ip_i$，即等于随机变量的取值乘以取值概率之和。现设$\xi$为连续型随机变量，其值域为区间$[a,b]$，分布密度为$p(x)$。我们先把整个区间分成长为 dx 的n个小区间，设其分点为$a=x_0,x_1,\cdots,x_n=b$考虑新随机变量ξ_n：

$$\begin{pmatrix} x_0, & x_1, & \cdots, & x_{n-1} \\ P(x_0\leqslant\xi<x_1), & P(x_1\leqslant\xi<x_2), & \cdots, & P(x_{n-1}\leqslant\xi<x_n) \end{pmatrix}$$

ξ_n是把ξ取值于小区间的概率集中在小区间的左端点上，显然分法越细，离散随机变量ξ_n就与连续型随机变量ξ越接近。随机变量ξ_n

的均值为

$$E\xi_n=\sum_{i=1}^{n}x_{i-1}P(x_{i-1}\leqslant\xi<x_i)$$
$$\doteq\sum_{i=1}^{n}x_{i-1}p(x_{i-1})dx.$$

当区间分法无限加细，ξ_n的均值就应逼近ξ的均值，可是这时上式变成

$$\int_a^b xp(x)dx=\int_{-\infty}^{\infty}xp(x)dx$$

（此处当x在$[a,b]$之外$p(x)=0$）

因此，连续型随机变量的均值定义如下：

定义3.4.1　设连续型随机变量ξ的分布密度为$p(x)$，而且积分$\int_{-\infty}^{\infty}xp(x)dx$绝对收敛，则$\xi$的均值定义为

$$E\xi=\int_{-\infty}^{\infty}xp(x)dx$$

例3.4.1　设ξ在$[a,b]$上均匀分布，则

$$E\xi=\int_a^b x\frac{1}{b-a}dx=\frac{1}{b-a}\left(\frac{b^2-a^2}{2}\right)=\frac{a+b}{2}$$

例3.4.2　设ξ的分布密度为

$$p(x)=\begin{cases}\frac{1}{(1500)^2}x & 0\leqslant x\leqslant1500\\ \frac{-1}{(1500)^2}(x-3000) & 1500<x\leqslant3000\\ 0 & \text{其它}\end{cases}$$

则
$$E\xi=\int_{-\infty}^{\infty}xp(x)dx$$
$$=\frac{1}{(1500)^2}\left[\int_0^{1500}x^2dx-\int_{1500}^{3000}x(x-3000)dx\right]$$
$$=1500$$

例3.4.3　设煤中杂质的百分比是ξ，ξ可看成是具有密度函数为

$$p(x)=\frac{1}{4875}x^2,\quad 10<x<25$$

的随机变量，则

$$E\xi=\frac{1}{4875}\int_{10}^{25}x^3dx\doteq 19.5$$

关于随机变量函数的均值，我们有如下定理：

定理3.4.1 （1）若 ξ 为一连续型随机变量，其分布密度为 $p(x)$，$f(x)$ 为一连续函数，则新随机变量 $f(\xi)$ 的均值为

$$Ef(\xi)=\int_{-\infty}^{\infty}f(x)p(x)dx \tag{3.4.1}$$

（2）若 (ξ,η) 为二维随机变量，联合密度函数为 $p(x,y)$，$H(x,y)$ 为二元连续函数，则新随机变量 $\zeta=H(\xi,\eta)$ 的均值为

$$E\zeta=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}H(x,y)p(x,y)dxdy \tag{3.4.2}$$

（证明较繁从略）

连续型随机变量均值的性质与离散型情形完全一致。

性质 1　对任意常数 c 有

$$E(c\xi)=cE\xi$$

性质 2　对任意 n 个随机变量 $\xi_1,\xi_2,\cdots,\xi_n$，若相应的均值都存在，则

$$E(\xi_1+\xi_2+\cdots+\xi_n)=E\xi_1+E\xi_2+\cdots+E\xi_n$$

证明　只就 $n=2$ 的情形证明就可以了。设 (ξ_1,ξ_2) 的联合分布密度为 $p(x,y)$，则

$$\begin{aligned}E(\xi_1+\xi_2)&=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x+y)p(x,y)dxdy\\&=\int_{-\infty}^{\infty}x\int_{-\infty}^{\infty}p(x,y)dydx\\&\quad+\int_{-\infty}^{\infty}y\int_{-\infty}^{\infty}p(x,y)dxdy\\&=\int_{-\infty}^{\infty}xp_1(x)dx+\int_{-\infty}^{\infty}yp_2(y)dy\\&=E\xi_1+E\xi_2\end{aligned}$$

性质 3　假设随机变量 ξ 与 η 相互独立，且均值都存在，则

$$E(\xi\eta)=E\xi\cdot E\eta$$

证明　设 (ξ,η) 的联合密度为$p(x,y)$，由独立性知 $p(x,y)=p_1(x)\cdot p_2(y)$，其中$p_1(x)$与$p_2(y)$分别是 ξ 与 η 的边缘分布密度，则

$$
\begin{aligned}
E\xi\eta &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xyp(x,y)dxdy \\
&= \int_{-\infty}^{\infty} xp_1(x)dx\cdot\int_{-\infty}^{\infty} yp_2(y)dy \\
&= E\xi\cdot E\eta
\end{aligned}
$$

例3.4.4　在很多问题里我们感兴趣的往往只是随机变量的大小而不注意其符号。就是说我们关心的是随机变量$|\xi|$。假如 ξ 的分布密度为

$$
p(x)=\begin{cases}\dfrac{e^{x}}{2} & \text{当}x\leqslant 0\text{时}\\[2mm] \dfrac{e^{-x}}{2} & \text{当}x>0\text{时}\end{cases}
$$

则有

$$
\begin{aligned}
E(|\xi|) &= \int_{-\infty}^{\infty}|x|p(x)dx \\
&= \frac{1}{2}\left[\int_{-\infty}^{0}(-x)e^{x}dx+\int_{0}^{\infty}xe^{-x}dx\right] \\
&= \frac{1}{2}[1+1]=1
\end{aligned}
$$

（二）方　　差

连续型随机变量的方差完全类似于离散情况定义为

$$D\xi=E(\xi-E\xi)^2$$

由(3.4.1)　我们有

$$D\xi=\int_{-\infty}^{\infty}(x-E\xi)^2p(x)dx$$

同离散情况一样的道理，$D\xi$可以简化计算为

$$D\xi=E[\xi^2]-[E\xi]^2$$

例3.4.5　设 ξ 在$[a,b]$上均匀分布，求$D\xi$。

先计算 $E\xi^2=\int_a^b x^2\frac{1}{b-a}dx=\frac{b^3-a^3}{3(b-a)}$

所以 $D\xi=E\xi^2-(E\xi)^2=\frac{b^3-a^3}{3(b-a)}-\left(\frac{a+b}{2}\right)^2$

$=\frac{(b-a)^2}{12}$

例3.4.6 设ξ的分布密度为

$$p(x)=\begin{cases}1+x & -1\leqslant x\leqslant 0\\ 1-x & 0\leqslant x\leqslant 1\end{cases}$$

(见图(3.4.1)). 由于ξ的密度函数的对称性所以显然$E\xi=0$

$$E\xi^2=\int_{-1}^{0}x^2(1+x)dx+\int_0^1 x^2(1-x)dx=\frac{1}{6}$$

所以 $D\xi=\frac{1}{6}$

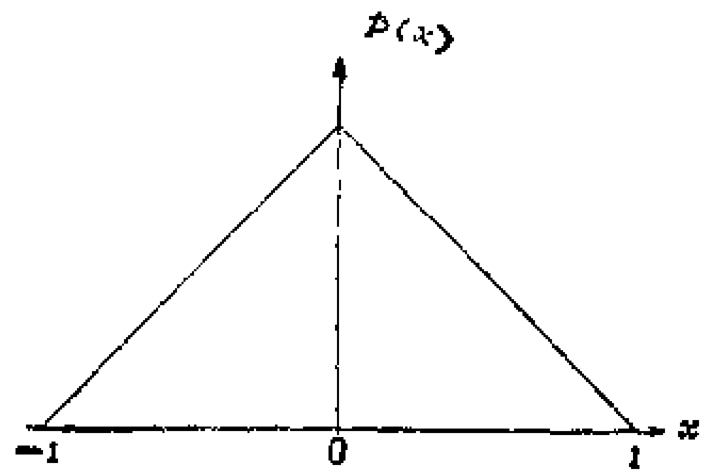


图 (3.4.1)

方差的性质与离散情况完全相同. 因这些性质的证明中只涉及到均值的性质，而均值的性质在连续情形与离散情形完全一样，所以离散情况下方差各性质的证明实质上对连续情形也适用. 所以下面我们只列出这些性质不加证明，证明请看第二章。

性质 1 对任一常数 c，有

$D(c\xi)=c^2D\xi$

性质 2 对 n 个相互独立随机变量$\xi_1,\xi_2,\cdots,\xi_n$，若每个方差都存在，则

$$D(\xi_1+\xi_2+\cdots+\xi_n)=D\xi_1+D\xi_2+\cdots+D\xi_n$$

性质 3 若$c\neq E\xi$，则$D\xi<E(\xi-c)^2$

关于相关系数及矩的概念和性质都与离散情况完全一样，只是

在计算时把求和换成求积分就是了。这里不再重复。

§3.5 几个重要的连续型随机变量

下面几个连续型随机变量不论在理论方面还是在应用方面，都是十分重要的随机变量。本节我们只介绍这些随机变量的定义及其基本性质，至于它们的应用将会在后面的章节里陆续看到。

(一) 正态分布

1 正态分布的定义

定义3.5.1 如果随机变量具有如下形式的分布密度：

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2\right)\qquad -\infty<x<\infty$$

则称它为服从正态分布的随机变量。

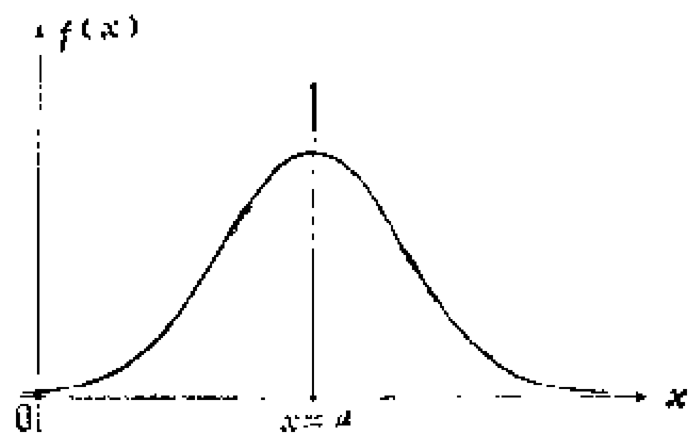


图 (3.5.1)

从密度函数$f(x)$的结构看，它是包含着两个参数μ和σ的一指数函数，其形状如'古钟'，在$x=\mu$点取极大值，并关于$x=\mu$成对称，而当x愈远离μ时$f(x)$值愈小（见图3.5.1）。一句话，$f(x)$是“中间高，两边低、呈对称”的曲线。以后为了方便起见，有时把具有参数μ、σ的正态分布简记为$N(\mu,\sigma^2)$，把随机变量服从参数为μ，σ的正态分布简记为$\xi\sim N(\mu,\sigma^2)$。

2 正态分布随机变量的线性函数 正态分布的一个重要性质是正态分布随机变量的线性函数仍然服从正态分布。我们通过下面两个定理来说明这个事实。

定理3.5.1 若随机变量$\xi\sim N(\mu,\sigma^2)$，则$\eta=\alpha\xi+\beta$也服从正态

分布，其中$\alpha\neq 0$，β为常数。

证明　设ξ的分布密度为$f(x)$，η的分布密度为$g(y)$，用公式(3.3.1)

$$g(y)=f(x)\left|\frac{dx}{dy}\right|$$

求$g(y)$。

由于这里$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$，而$x=\frac{y-\beta}{\alpha}$，$\left|\frac{dx}{dy}\right|=\frac{1}{|\alpha|}$，

所以

$$g(y)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{\left(\frac{y-\beta}{\alpha}-\mu\right)^2}{2\sigma^2}}\cdot\frac{1}{|\alpha|}$$

$$=\frac{1}{|\alpha|\sigma\sqrt{2\pi}}e^{-\frac{[y-(\alpha\mu+\beta)]^2}{(2\alpha\sigma)^2}}$$

即$\eta\sim N((\alpha\mu+\beta),(\alpha\sigma)^2)$

定理3.5.2　若$\xi_1,\xi_2,\cdots,\xi_n$为n个相互独立的服从正态分布的随机变量，则它们的和$\sum_{i=1}^{n}\xi_i$也服从正态分布。

证明　只就$n=2$证明就可以了.设$\xi_1\sim N(\mu_1,\sigma_1^2)$，$\xi_2\sim N(\mu_2,\sigma_2^2)$，并设$p_1(x)$，$p_2(x)$，$g(x)$分别为$\xi_1$，$\xi_2$及$\xi_1+\xi_2$的分布密度，则根据(3.3.8)有

$$g(x)=\int_{-\infty}^{\infty}p_1(x-z)p_2(z)dz$$

$$=\frac{1}{\sigma_1\sigma_2(2\pi)}\int_{-\infty}^{\infty}e^{-\frac{1}{2}\left[\frac{(x-z-\mu_1)^2}{\sigma_1^2}+\frac{(z-\mu_2)^2}{\sigma_2^2}\right]}dz$$

为简单起见，令$u=z-\mu_2$，$y=x-\mu_1-\mu_2$，则上式成为

$$\frac{1}{\sigma_1\sigma_2(2\pi)}\int_{-\infty}^{\infty}e^{-\frac{1}{2}\left[\frac{(y-u)^2}{\sigma_1^2}+\frac{u^2}{\sigma_2^2}\right]}du \qquad (3.5.1)$$

计算此积分，先分析

$$\frac{(y-u)^2}{\sigma_1^2}+\frac{u^2}{\sigma_2^2}=\frac{y^2+u^2-2uy}{\sigma_1^2}+\frac{u^2}{\sigma_2^2}$$

$$=\frac{u^2(\sigma_1^2+\sigma_2^2)}{\sigma_1^2\sigma_2^2}-2u\frac{y}{\sigma_1^2}+\frac{y^2}{\sigma_1^2}$$

$$=\left[\frac{u\sqrt{\sigma_1^2+\sigma_2^2}}{\sigma_1\sigma_2}-\frac{y}{\sigma_1}\cdot\frac{\sigma_2}{\sqrt{\sigma_1^2+\sigma_2^2}}\right]^2$$

$$+\frac{y^2}{\sigma_1^2}\left[1-\frac{\sigma_2^2}{\sigma_1^2+\sigma_2^2}\right]$$

$$=\left[\frac{u\sqrt{\sigma_1^2+\sigma_2^2}}{\sigma_1\sigma_2}-\frac{\sigma_2 y}{\sigma_1\sqrt{\sigma_1^2+\sigma_2^2}}\right]^2+\frac{y^2}{\sigma_1^2+\sigma_2^2}$$

作变量替换

$$t=\left[\frac{u\sqrt{\sigma_1^2+\sigma_2^2}}{\sigma_1\sigma_2}-\frac{\sigma_2 y}{\sigma_1\sqrt{\sigma_1^2+\sigma_2^2}}\right]\qquad dt=\frac{\sqrt{\sigma_1^2+\sigma_2^2}}{\sigma_1\sigma_2}du$$

代入(3.5.1)则得

$$g(x)=\frac{1}{\sqrt{\sigma_1^2+\sigma_2^2}\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{y^2}{\sigma_1^2+\sigma_2^2}\right)}$$

$$\cdot\int_{-\infty}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}dt$$

$$=\frac{1}{\sqrt{\sigma_1^2+\sigma_2^2}\sqrt{2\pi}}e^{-\frac{1}{2}\frac{(x-(\mu_1+\mu_2))^2}{\sigma_1^2+\sigma_2^2}}$$

所以$\xi_1+\xi_2\sim N(\mu_1+\mu_2,\sigma_1^2+\sigma_2^2)$

例3.5.1　设$\xi\sim N(\mu,\sigma^2)$由定理3.5.1知

$$\eta=\frac{\xi-\mu}{\sigma}=\frac{1}{\sigma}\xi-\frac{\mu}{\sigma}\sim N(0,1)$$

（3）正态分布随机变量的均值和方差

定理3.5.3　设$\xi\sim N(\mu,\sigma^2)$，则$E\xi=\mu$，$D\xi=\sigma^2$。

证明　$E\xi=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}xe^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$，令 $z=(x-\mu)/\sigma$，这时

$dx=\sigma dz$，我们有

$$E\xi=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\sigma z+\mu)e^{-\frac{z^2}{2}}dz$$

$$=\frac{1}{\sqrt{2\pi}}\sigma\int_{-\infty}^{\infty}ze^{-\frac{z^2}{2}}dz+\mu\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{z^2}{2}}dz$$

上式中第一个积分的被积函数为奇函数，所以它在整个数轴上的积分值为零，故求得

$$E\xi=\mu$$

为求方差下面考虑 $E\xi^2=\frac{1}{\sqrt{2\pi}\sigma}\int_{-\infty}^{\infty}x^2e^{-\frac{(x-\mu)^2}{2\sigma^2}}dx$

还令 $z=(x-\mu)/\sigma$ 则

$$E\xi^2=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}(\sigma z+\mu)^2e^{-\frac{z^2}{2}}dz$$

$$=\frac{1}{\sqrt{2\pi}}\sigma^2\int_{-\infty}^{\infty}z^2e^{-z^2/2}dz+2\mu\sigma\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}ze^{-z^2/2}dz$$

$$+\mu^2\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{z^2}{2}}dz$$

上式中第二个积分值还是零，第三项为 μ^2，所以关键是求出第一项积分。用分部积分法，令 $ze^{-\frac{z^2}{2}}dz=dv, z=u$，则 $v=-e^{-\frac{z^2}{2}}, dz=du$，于是积分

$$\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}z^2e^{-\frac{z^2}{2}}dz=-\frac{1}{\sqrt{2\pi}}ze^{-\frac{z^2}{2}}\Big|_{-\infty}^{\infty}$$

$$+\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{z^2}{2}}dz$$

$$=0+1=1$$

所以 $E\xi^2=\sigma^2+\mu^2$，故 $D\xi=E\xi^2-(E\xi)^2=\sigma^2$

通过求正态分布的均值和方差，我们也明确了正态分布密度函数中的两个参数 μ,σ 的意义，它们分别是正态分布随机变量的均值

和均方差．因此，我们可以得出这样的结论：正态分布随机变量的概率规律，由它的均值和方差完全确定．从其密度函数的形状上看，它是关于均值对称的“钟”形曲线，而且其“胖瘦”情形由σ大小决定，σ较大时，密度曲线显得‘平些’，σ越小，图形越陡（见图3.5.2）

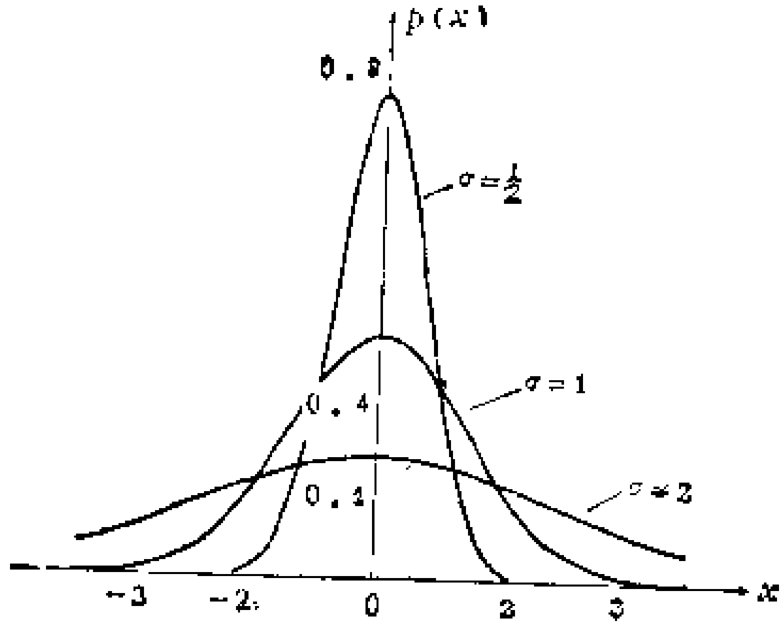


图 (3.5.2)

正态分布随机变量取值集中于均值附近的程度可用其均方差σ为尺度衡量，若$\xi\sim N(\xi,\sigma^2)$则

1．$P\{\mu-\sigma\leqslant\xi\leqslant\mu+\sigma\}=68\%$

2．$P\{\mu-2\sigma\leqslant\xi\leqslant\mu+2\sigma\}=95.5\%$

3．$P\{\mu-3\sigma\leqslant\xi\leqslant\mu+3\sigma\}=99.7\%$（见图3.5.3）

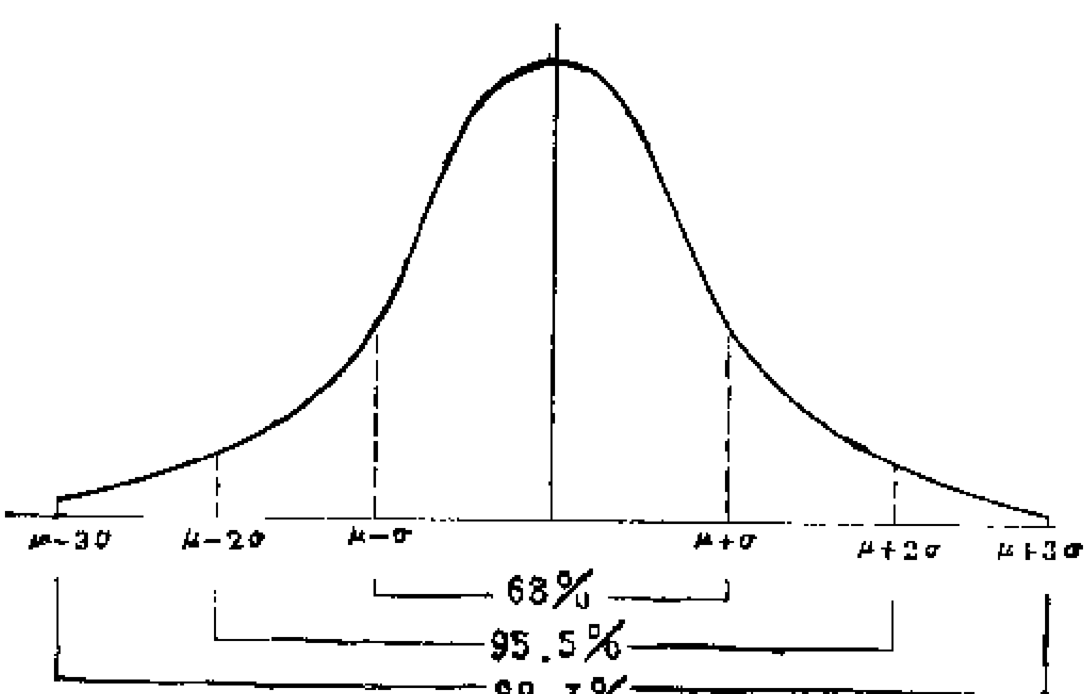


图 (3.5.3)

在实际应用中下面公式也是有用的

1. $P\{\mu-1.96\sigma\leqslant\xi\leqslant\mu+1.96\sigma\}=95\%$

2. $P\{\mu-2.58\sigma\leqslant\xi\leqslant\mu+2.58\sigma\}=99\%$

3. $P\{\mu-3.29\sigma\leqslant\xi\leqslant\mu+3.29\sigma\}=99.9\%$

欲求 ξ 取值于任一区间 $[a,b]$ 的概率，均可通过查正态表得到。正态表是对 $N(0,1)$ 编制的，其中有密度函数 $\phi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ 的数值表，也有分布函数 $\Phi(x)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{u^2}{2}}du$ 的数值表。

如求概率 $P(a\leqslant\xi\leqslant b)$ 可通过查 $\Phi(x)$ 表得到。事实上，若 $\xi\sim N(\mu,\sigma^2)$ 则其标准化 $\eta=\frac{\xi-\mu}{\sigma}\sim N(0,1)$ 所以

$$
\begin{aligned}
P(a\leqslant\xi\leqslant b)&=P\left\{\frac{a-\mu}{\sigma}\leqslant\frac{\xi-\mu}{\sigma}\leqslant\frac{b-\mu}{\sigma}\right\}\\
&=\frac{1}{\sqrt{2\pi}}\int_{\frac{a-\mu}{\sigma}}^{\frac{b-\mu}{\sigma}}e^{-\frac{u^2}{2}}du\\
&=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{b-\mu}{\sigma}}e^{-\frac{u^2}{2}}du-\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\frac{a-\mu}{\sigma}}e^{-\frac{u^2}{2}}du\\
&=\Phi\left(\frac{b-\mu}{\sigma}\right)-\Phi\left(\frac{a-\mu}{\sigma}\right)
\end{aligned}
$$

在查 $\Phi(x)$ 表时，注意它是只对 x 是正值编制的，对负 x 值，可通过 $\Phi(-x)+\Phi(x)=1$ 来确定。

（二）二维正态分布

定义3.5.2　二维随机变量 (ξ,η) 当它的联合分布密度具有如下形式时称其为服从二维正态分布。

$$f(x,y)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}e^{-Q(x,y)/2},\tag{3.5.2}$$

其中 $\sigma_1>0,\sigma_2>0,|\rho|<1$，而 $Q(x,y)$ 是下面的二次型

$$Q(x,y)=\frac{1}{1-\rho^2}\left[\left(\frac{x-\mu_1}{\sigma_1}\right)^2-2\rho\left(\frac{x-\mu_1}{\sigma_1}\right)\left(\frac{y-\mu_2}{\sigma_2}\right)\right.$$

$$\left.+\left(\frac{y-\mu_2}{\sigma_2}\right)^2\right]$$

μ_1,μ_2均为常数．$f(x,y)$包含五个参数：$\mu_1,\mu_2,\sigma_1^2,\sigma_2^2,\rho$。

首先应验证（3.5.2）式定义的二元函数 $f(x,y)$是一个密度函数．事实上它是下述定理的一个直接结果．

定理3.5.4　定义3.5.2中定义的函数$f(x,y)$是一个二元密度函数，其边缘分布分别是$N(\mu_1,\sigma_1^2)$，$N(\mu_2,\sigma_2^2)$，而 ρ 是 ξ 和 η 的相关系数．

证明　令 $f_1(x)=\int_{-\infty}^{\infty}f(x,y)dy$．又注意到

$$Q(x,y)=\frac{1}{1-\rho^2}\left(\frac{y-\mu_2}{\sigma_2}-\rho\frac{x-\mu_1}{\sigma_1}\right)^2+\left(\frac{x-\mu_1}{\sigma_1}\right)^2$$

$$=\frac{1}{1-\rho^2}\left\{\frac{y-[\mu_2+\rho(\sigma_2/\sigma_1)(x-\mu_1)]}{\sigma_2}\right\}^2$$

$$+\left(\frac{x-\mu_1}{\sigma_1}\right)^2$$

可推出

$$f_1(x)=\frac{1}{\sigma_1\sqrt{2\pi}}exp\left[\frac{-(x-\mu_1)^2}{2\sigma_1^2}\right]\cdot\int_{-\infty}^{\infty}\frac{1}{\sigma_2\sqrt{2\pi}\sqrt{1-\rho^2}}$$

$$exp\{-(y-\beta_x)^2/[2\sigma_2^2(1-\rho^2)]\}dy \qquad (3.5.3)$$

其中　$\beta_x=\mu_2+\rho(\sigma_2/\sigma_1)(x-\mu_1)$。

（3.5.3）式中的被积函数恰是 $N(\beta_x,\sigma_2^2(1-\rho^2))$ 随机变量的密度函数，故积分为 1．因此，

$$f_1(x)=\frac{1}{\sigma_1\sqrt{2\pi}}exp\left\{\frac{-(x-\mu_1)^2}{2\sigma_1^2}\right\} \qquad -\infty<x<\infty.$$

由此立刻推出

$$\int_{-\infty}^{\infty}\left\{\int_{-\infty}^{\infty}f(x,y)dy\right\}dx=\int_{-\infty}^{\infty}f_1(x)dx=1$$

即证得$f(x,y)$是二维随机变量的密度函数，ξ 的边缘分布是$N(\mu_1,\sigma_1^2)$。类似可证 η 的边缘分布为$N(\mu_2,\sigma_2^2)$。

下面计算ξ,η的相关系数，为此先求

$$E(\xi-\mu_1)(\eta-\mu_2)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)f(x,y)dxdy$$

$$=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}}\int_{-\infty}^{\infty}e^{-\frac{(y-\mu_2)^2}{2\sigma_2^2}}dy$$

$$\cdot\int_{-\infty}^{\infty}(x-\mu_1)(y-\mu_2)$$

$$\cdot exp\left\{-\frac{1}{2(1-\rho^2)}\left(\frac{x-\mu_1}{\sigma_1}\right.\right.$$

$$\left.\left.-\rho\frac{y-\mu_2}{\sigma_2}\right)^2\right\}dx$$

作变换 $z=\frac{1}{\sqrt{1-\rho^2}}\left(\frac{x-\mu_1}{\sigma_1}-\rho\frac{y-\mu_2}{\sigma_2}\right)$，$w=\frac{y-\mu_2}{\sigma_2}$，上式化为

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(\sigma_1\sigma_2\sqrt{1-\rho^2}wz+\rho\sigma_1\sigma_2w^2)e^{-\frac{w^2}{2}-\frac{z^2}{2}}dzdw$$

$$=\frac{\rho\sigma_1\sigma_2}{2\pi}\int_{-\infty}^{\infty}w^2e^{-\frac{w^2}{2}}dw\int_{-\infty}^{\infty}e^{-\frac{z^2}{2}}dz$$

$$+\frac{\sigma_1\sigma_2\sqrt{1-\rho^2}}{2\pi}\int_{-\infty}^{\infty}we^{-\frac{w^2}{2}}dw\int_{-\infty}^{\infty}ze^{-\frac{z^2}{2}}dz$$

上式中第一、第二个积分都等于$\sqrt{2\pi}$，第三、第四个积分都为0，故

$$E(\xi-\mu_1)(\eta-\mu_2)=\rho\sigma_1\sigma_2$$

这样，ξ,η的相关系数

$$\frac{E(\xi-\mu_1)(\eta-\mu_2)}{\sqrt{D\xi}\sqrt{D\eta}}=\frac{\rho\sigma_1\sigma_2}{\sigma_1\sigma_2}=\rho$$

定理3.5.5 若(ξ,η)是二维正态随机变量，则ξ与η相互独立的充分必要条件是其相关系数$\rho=0$。

证明 必要性。若ξ与η相互独立，即对一切x,y有

$$f(x,y)=f_1(x)\cdot f_2(y)$$

其中$f(x,y)$，$f_1(x)$和$f_2(y)$含义同定理3.5.4。

特别有 $f(\mu_1,\mu_2)=f_1(\mu_1)\cdot f_2(\mu_2)$，而

$$f(\mu_1,\mu_2)=\frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}},$$

$$f_1(\mu_1)=\frac{1}{\sqrt{2\pi}\,\sigma_1},\qquad f_2(\mu_2)=\frac{1}{\sqrt{2\pi}\,\sigma_2},$$

于是 $$\frac{1}{\sqrt{1-\rho^2}}=1$$

从而 $\rho=0$。

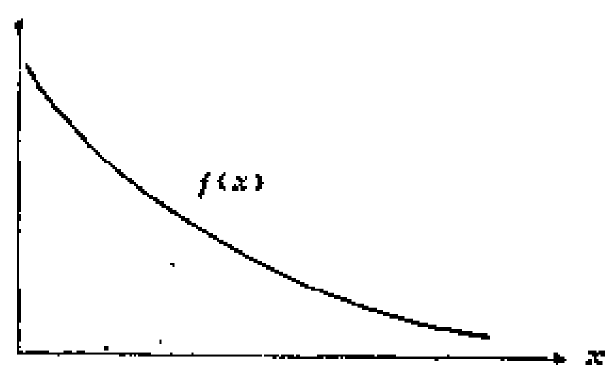
充分性的证明显然。

（三）指数分布和加马（Gamma）分布

定义3.5.3 取非负值的连续型随机变量ξ，若其分布密度为

$$f(x)=\begin{cases}\alpha e^{-\alpha x} & x>0\\ 0 & 其它\end{cases}$$

其中 $\alpha>0$（见图3.5.4），则称ξ服从参数为α的指数分布



图（3.5.4）

指数分布在可靠性理论里起着重要的作用，它是一个很重要的分布.

在实际中有这样一类元件，如保险丝、宝石轴承、玻璃制品等，它们的共同特点是不存在"磨损"而变旧的问题．接在电路里的保险丝不能因为已经使用了一段时间而被认为快要熔断了；镶在窗户上的玻璃不能因为已经使用了一段时间而被认为快要破了；宝石轴承耐磨程度之大使人可以认为它没有被磨损的可能．换句话说，这种元件损坏与否与过去使用历史无关，如果在时刻t元件没有损坏，那么，在t以后继续使用就应看成和使用新的一样，可以验证，上述元件的寿命就是服从指数分布的随机变量．这里不作详细介绍了。

定理3.5.6　设ξ服从参数为α的指数分布，则$E\xi=\frac{1}{\alpha}$，$D\xi=\frac{1}{\alpha^2}$，其分布函数$F(x)=1-e^{-\alpha x}(x\geqslant 0)$

证明　$F(x)=P(\xi\leqslant x)=\int_0^x \alpha e^{-\alpha t}\,dt=\begin{cases}1-e^{-\alpha x} & x>0\\ 0 & \text{其它}\end{cases}$

$$E\xi=\int_0^\infty x\alpha e^{-\alpha x}dx$$

利用分部积分法，令$\alpha e^{-\alpha x}dx=dv, x=u$，则$v=-e^{-\alpha x}$，$du=dx$，所以

$$E\xi=[-xe^{-\alpha x}]_0^\infty+\int_0^\infty e^{-\alpha x}dx=\frac{1}{\alpha}$$

类似方法可求出　$E\xi^2=\frac{2}{\alpha^2}$

所以　$D\xi=(E\xi^2)-(E\xi)^2=\frac{1}{\alpha^2}$

从指数分布的密度函数结构可以看出，它是只包含一个参数的分布。而定理3.5.6告诉我们，这个参数正好是其均值的倒数，因此对指数分布来说，随机变量的均值就完全确定了其整个的取值规律。

例3.5.2　设ξ服从参数为α的指数分布，求概率$P(\xi>E\xi)$。

因为　$E\xi=\frac{1}{\alpha}$，所以

$$P(\xi>E\xi)=P(\xi>\frac{1}{\alpha})=1-F\left(\frac{1}{\alpha}\right)=e^{-\alpha(\frac{1}{\alpha})}$$

$$=e^{-1}<\frac{1}{2}$$

下面要讨论的 Gamma 分布是指数分布的自然推广。为了便于叙述，先介绍 Gamma 函数和贝他（*Beta*）函数的概念。

定义3.5.4　函数　$\Gamma(p)=\int_0^\infty x^{p-1}e^{-x}dx$，$p>0$　称为Gamma函数。

把$\Gamma(p)$用分部积分法积分，令$e^{-x}dx = dv, x^{p-1} = u$，则

$$\Gamma(p) = -e^{-x}x^{p-1}\Big|_0^{\infty} - \int_0^{\infty}[-e^{-x}(p-1)x^{p-2}]dx$$
$$= 0 + (p-1)\int_0^{\infty}e^{-x}x^{p-2}dx$$
$$= (p-1)\Gamma(p-1)$$

这说明Gamma 函数有递推关系。若 p 为正整数，即$p = n$时，$\Gamma(n) = (n-1)\Gamma(n-1) = (n-1)(n-2)\Gamma(n-2) = (n-1)(n-2)\cdots\Gamma(1)$，而 $\Gamma(1) = \int_0^{\infty}e^{-x}dx = 1$，因此$\Gamma(n) = (n-1)!$ ，所以Gamma函数可看成阶乘函数的推广。

对于Gamma函数$\Gamma(p)$来说，当 $p = \frac{1}{2}$时的函数值很重要。可以证明 $\Gamma(\frac{1}{2}) = \sqrt{\pi}$。事实上

$$\Gamma\left(\frac{1}{2}\right) = 2\int_0^{\infty}e^{-x^2}dx \text{或} \Gamma\left(\frac{1}{2}\right) = 2\int_0^{\infty}e^{-y^2}dy$$

所以

$$\left[\Gamma\left(\frac{1}{2}\right)\right]^2 = \left(2\int_0^{\infty}e^{-x^2}dx\right)\left(2\int_0^{\infty}e^{-y^2}dy\right)$$
$$= 4\int_0^{\infty}dx\int_0^{\infty}e^{-(x^2+y^2)}dy$$

利用极坐标变换，则

$$\left[\Gamma\left(\frac{1}{2}\right)\right]^2 = 4\int_0^{\frac{\pi}{2}}d\theta\int_0^{\infty}e^{-r^2}rdr$$

$$= 4\frac{\pi}{2}\left[\frac{e^{-r^2}}{-2}\right]_0^{\infty} = \pi$$

所以 $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$。

定义3.5.5 设p,q为正实数，函数

$$B(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1}dx$$

称为 Beta 函数。

可以证明 Gamma函数与 Bata 函数间存在下面重要关系：

$$B(p,q)=\frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}^{*} \tag{3.5.4}$$

下面借助Gamma函数来定义 Gamma分布。

定义3.5.6 设ξ为取非负值的连续型随机变量，若其分布密度为

$$f(x)=\begin{cases}\frac{\alpha}{\Gamma(\gamma)}(\alpha x)^{\gamma-1}e^{-\alpha x} & x>0\\ 0 & 其它\end{cases} \tag{3.5.5}$$

其中$\gamma\geqslant 0,\alpha>0$，则称$\xi$为服从Gamma分布的随机变量。显然 Gamma分布中具有两个参数：α和γ，γ称为形状参数，对不同的γ值Γ曲线的形状是不同的，图(3.5.5)是当$\alpha=1$时，不同γ值的Γ分布曲线的形状。ξ服从参数为α,γ的 Gamma 分布简记为$\xi\sim G(\alpha,\gamma)$。

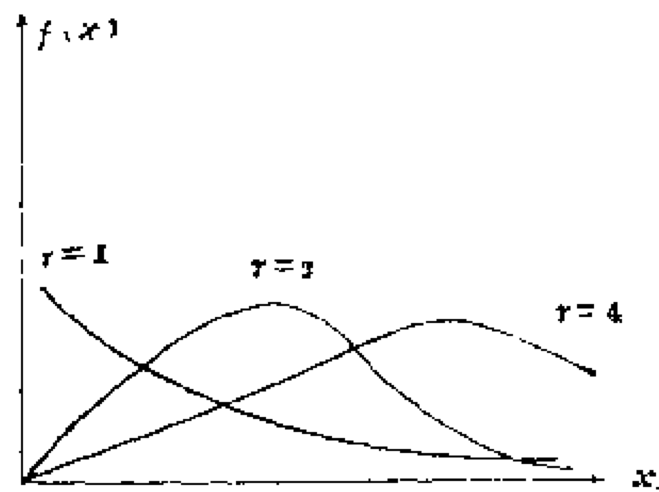


图 (3.5.5)

当$\gamma=1$时 Gamma 分布密度函数变成$f(x)=\alpha e^{-\alpha x}$，所以指数分布是Gamma分布的特殊情形。当$\gamma$为正整数时的 Gamma 分布称为 Erlangian 分布。当$\alpha=\frac{1}{2}$，$\gamma=\frac{n}{2}$，n为正整数时，Gamma 分布就是在以后数理统计里常要使用的χ^2分布。

定理3.5.7 若ξ为服从参数α和γ的 Gamma 分布的随机变量，则

$$E\xi=\frac{\gamma}{\alpha},\ D\xi=\frac{\gamma}{\alpha^2}.$$

---

* 参看克拉美《统计学数学方法》

证明 $E\xi = \int_0^{\infty} x \frac{\alpha}{\Gamma(\gamma)}(\alpha x)^{\gamma-1} e^{-\alpha x} dx$

$$= \int_0^{\infty} \frac{(\alpha x)^{\gamma} e^{-\alpha x}}{\Gamma(\gamma)} dx \qquad 令 t = \alpha x$$

$$= \frac{1}{\alpha \Gamma(\gamma)} \int_0^{\infty} t^{\gamma} e^{-t} dt = \frac{\Gamma(\gamma+1)}{\alpha \Gamma(\gamma)} = \frac{\gamma}{\alpha}$$

$$E\xi^2 = \int_0^{\infty} x^2 \frac{\alpha^{\gamma} x^{\gamma-1} e^{-\alpha x}}{\Gamma(\gamma)} dx$$

$$= \int_0^{\infty} \frac{(\alpha x)^{\gamma+1} e^{-\alpha x}}{\alpha \Gamma(\gamma)} dx = \int_0^{\infty} \frac{t^{\gamma+1} e^{-t}}{\alpha^2 \Gamma(\gamma)} dt$$

$$= \frac{\Gamma(\gamma+1)}{\alpha^2 \Gamma(\gamma)} = \frac{(\gamma+1)\gamma}{\alpha^2}$$

所以 $D\xi = \frac{(\gamma+1)\gamma}{\alpha^2} - \left(\frac{\gamma}{\alpha}\right)^2 = \frac{\gamma}{\alpha^2}$。

定理3.5.8 若ξ,η相互独立，分别服从参数为(α,γ_1)和(α,γ_2)的Gamma分布，则其和$\xi+\eta$服从参数为$(\alpha,\gamma_1+\gamma_2)$的Gamma分布。

证明 设ξ,η的分布密度分别记为f_ξ和f_η，则$\xi+\eta$的密度函数

$$q(x) = \int_{-\infty}^{\infty} f_\xi(z) f_\eta(x-z) dz$$

$$= \int_0^{x} \frac{\alpha}{\Gamma(\gamma_1)}(\alpha z)^{\gamma_1-1} e^{-\alpha z} \frac{\alpha}{\Gamma(\gamma_2)}[\alpha(x-z)]^{\gamma_2-1} e^{-\alpha(x-z)} dz$$

$$= \frac{\alpha^{\gamma_1+\gamma_2} e^{-\alpha x}}{\Gamma(\gamma_1)\Gamma(\gamma_2)} \int_0^{x} (x-z)^{\gamma_2-1} z^{\gamma_1-1} dz, \quad x > 0.$$

令 $u = \frac{z}{x}$，则

$$q(x) = \frac{\alpha^{\gamma_1+\gamma_2}}{\Gamma(\gamma_1)\Gamma(\gamma_2)} e^{-\alpha x} x^{\gamma_1+\gamma_2-1} \int_0^1 u^{\gamma_1-1}(1-u)^{\gamma_2-1} du$$

$$= \frac{\alpha^{\gamma_1+\gamma_2}}{\Gamma(\gamma_1)\Gamma(\gamma_2)} e^{-\alpha x} x^{\gamma_1+\gamma_2-1} \frac{\Gamma(\gamma_1)\Gamma(\gamma_2)}{\Gamma(\gamma_1+\gamma_2)}$$

$$= \frac{\alpha}{\Gamma(\gamma_1+\gamma_2)} (\alpha x)^{\gamma_1+\gamma_2-1} e^{-\alpha x}$$

（四）数理统计里常用的三个分布

1 χ^2分布 χ^2分布是Gamma分布的一种特殊情形。在 Gamma 分布中，令 $\alpha=\frac{1}{2},\gamma=\frac{n}{2}$，其中 n 是正整数，就得到自由度为 n 的 χ^2分布，其分布密度为

$$f(x)=\begin{cases}\dfrac{1}{\Gamma\left(\dfrac{n}{2}\right)2^{\frac{n}{2}}}x^{\frac{n}{2}-1}e^{-\frac{x}{2}}, & x>0\\ 0 & x\leqslant 0\end{cases}\qquad(3.5.6)$$

对不同的自由度n，$f(x)$的图形如图（3.5.6），随机变量ξ服从自由度为 n 的 χ^2分布简记为 $\xi\sim\chi^2(n)$。

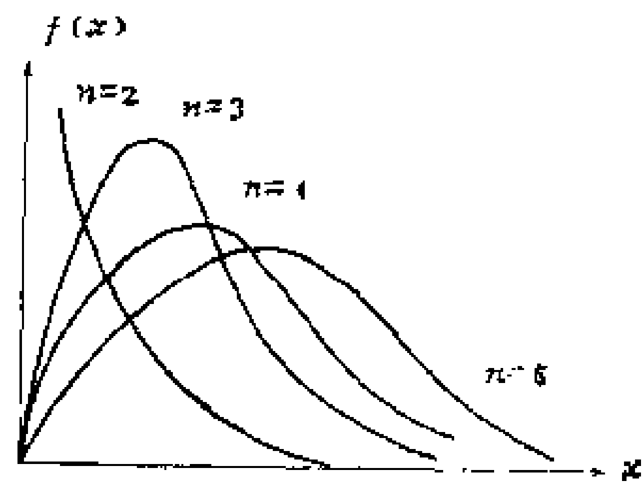


图 （3.5.6）

χ^2分布在数理统计里是十分重要的，对应用来说下一个定理是基础。

定理3.5.9 若$\xi_1,\xi_2,\cdots,\xi_n$是相互独立同 N(0,1) 分布的随机变量，则随机变量

$$\chi^2=\xi_1^2+\xi_2^2+\cdots+\xi_n^2$$

服从自由度为 n 的χ^2分布。

证明 首先$\xi_i^2\sim\chi^2(1)$，$i=1,2,\cdots,n$。这是因为

$$P(\xi_i^2\leqslant y)=P(-\sqrt{y}\leqslant\xi_i\leqslant\sqrt{y})$$
$$=\Phi(\sqrt{y})-\Phi(-\sqrt{y})$$

这里$\Phi(\cdot)$表示N(0,1)随机变量的分布函数。由此得到 ξ_i^2的分布密度

$$q(y)=\begin{cases}\dfrac{1}{\sqrt{2\pi y}}e^{-y/2} & y>0\\ 0 & y\leqslant 0\end{cases}$$

其次，由定理3.5.8知，若两个随机变量 ξ，η 相互独立，$\xi\sim\chi^2(n)$，$\eta\sim\chi^2(m)$，则$\xi+\eta\sim\chi^2(n+m)$.因此 n 个相互独立的 $\chi^2(1)$ 分布的随机变量之和

$$\chi^2=\xi_1^2+\xi_2^2+\cdots+\xi_n^2$$

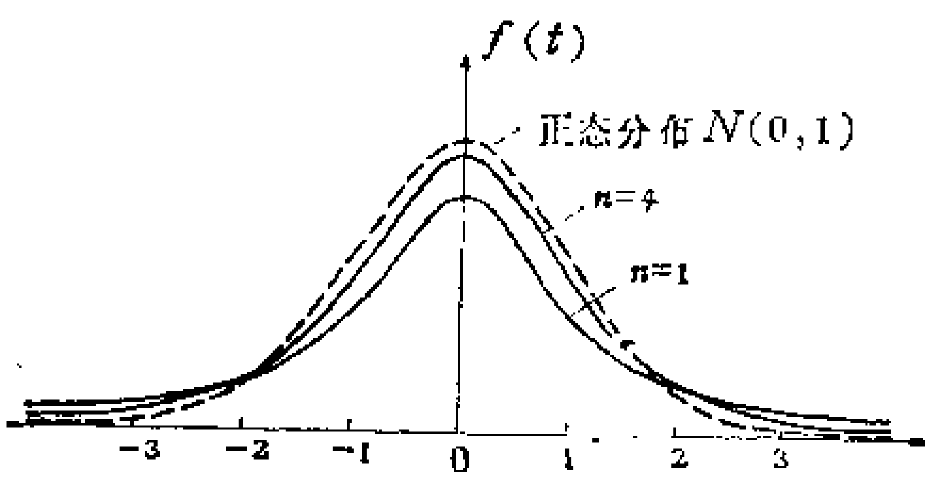
服从自由度为 n 的 χ^2分布。

2　t 分布

t 分布的分布密度定义为

$$f(t)=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-\frac{n+1}{2}}\quad(-\infty<t<\infty)\quad(3.5.7)$$

n 称为分布的自由度。由于 $f(t)=f(-t)$，所以它的图形关于纵轴对称，当 n 增大时，它近似于正态分布，如图（3.5.7）。



t 分布

图（3.5.7）

定理3.5.10　若随机变量X和Y独立，X服从正态分布 N(0，1)，Y服从自由度为 n 的χ^2分布，则

$$T=X\Big/\sqrt{\frac{Y}{n}}$$

服从自由度为 n 的 t 分布。简记为$T\sim t(n)$。

证明　设Y的分布密度为 $f_Y(y)$，$Z=\sqrt{\dfrac{Y}{n}}$的分布密度为 $f_Z(z)$，则

$$f_Z(z)=f_Y(y)\left|\frac{dy}{dz}\right|=f_Y(nz^2)\left|\frac{dy}{dz}\right|$$

$$=\frac{1}{\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}}}n^{\frac{n}{2}-1}z^{2\left(\frac{n}{2}-1\right)}e^{-\frac{nz^2}{2}}2nz$$

$$=\frac{1}{\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}-1}}n^{\frac{n}{2}}z^{n-1}e^{-\frac{nz^2}{2}}\quad(z>0)$$

根据求随机变量商的分布的公式，得 $T=X\Big/\sqrt{\frac{Y}{n}}$ 的分布密度为

$$f(t)=\int_0^{\infty}\frac{e^{-(ty)^2/2}}{\sqrt{2\pi}}\cdot\frac{e^{-\frac{ny^2}{2}}y^{n-1}n^{\frac{n}{2}}}{2^{\frac{n}{2}-1}\Gamma\left(\frac{n}{2}\right)}y\,dy$$

$$=\int_0^{\infty}\frac{e^{-y^2(t^2+n)/2}(y\sqrt{n})^n}{2^{(n-1)/2}\sqrt{\pi}\,\Gamma\left(\frac{n}{2}\right)}dy$$

作变量变换$x=y^2(n+t^2)/2$，则得

$$f(t)=\int_0^{\infty}\frac{e^{-x}x^{(n-1)/2}}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)\left(1+\frac{t^2}{n}\right)^{(n+1)/2}}dx$$

$$=\frac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-(n+1)/2}\quad(-\infty<t<\infty)$$

或者$f(t)$可写成

$$f(t)=\frac{1}{\sqrt{n}\,B\left(\frac{1}{2},\frac{n}{2}\right)}\left(1+\frac{t^2}{n}\right)^{-(n+1)/2}$$

（三） F 分布。

定理3.5.11 若ξ,η独立，分别服从自由度为 m,n 的χ^2分布，则随机变量

$$\mathrm{F}=\frac{\xi/m}{\eta/n}$$

的分布密度为

$$f_F(x)=\begin{cases}\dfrac{\left(\dfrac{m}{n}\right)^{m/2}x^{m/2-1}}{B\left(\dfrac{m}{2},\dfrac{n}{2}\right)}\left(1+\dfrac{m}{n}x\right)^{-\frac{m+n}{2}}, & x>0\\ 0 & x\leqslant 0\end{cases}\tag{3.5.8}$$

F 称为服从自由度为(m,n)的 F 分布，简记为 $F\sim F(m,n)$. 注意. F 分布具有两个自由度，第一个自由度m是组成 F 分布的分子随机变量的自由度，第二个自由度n 是分母的自由度.密度函数 $f_F(x)$的图形见图（3.5.8）.

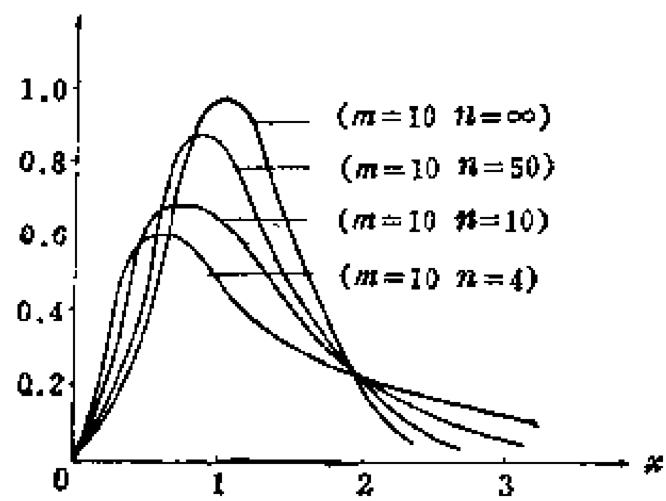


图 （3.5.8）

证明　由（3.5.7)可推知$\dfrac{\xi}{m}$的分布密度为

$$f_1(y)=\begin{cases}\dfrac{m^{\frac{m}{2}}}{2^{\frac{m}{2}}\Gamma\left(\dfrac{m}{2}\right)}y^{\frac{m}{2}-1}e^{-\frac{my}{2}} & (y>0)\\ 0 & (y\leqslant 0)\end{cases}$$

而$\dfrac{\eta}{n}$的分布密度为

$$f_2(y)=\begin{cases}\dfrac{n^{\frac{n}{2}}}{2^{\frac{n}{2}}\Gamma\left(\dfrac{n}{2}\right)}y^{\frac{n}{2}-1}e^{-\frac{ny}{2}} & (y>0)\\ 0 & (y\leqslant 0)\end{cases}$$

因此 $F=\dfrac{\xi/m}{\eta/n}$ 的分布密度为

$$f_F(x)=c\int_0^{\infty}y(xy)^{\frac{m}{2}-1}e^{-\frac{mxy}{2}}y^{\frac{n}{2}-1}e^{-\frac{ny}{2}}dy$$

$$=cx^{\frac{m}{2}-1}\int_0^{\infty}y^{\frac{m+n}{2}-1}e^{-\frac{1}{2}y(mx+n)}dy$$

其中 $$c=\frac{m^{\frac{m}{2}}n^{\frac{n}{2}}}{2^{\frac{m+n}{2}}\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)} \tag{3.5.9}$$

令$z=y(mx+n)$，则

$$f_F(x)=c\frac{x^{\frac{m}{2}-1}}{(mx+n)^{\frac{m+n}{2}}}\int_0^{\infty}z^{\frac{m+n}{2}-1}e^{-\frac{z}{2}}dz$$

$$=c\frac{x^{\frac{m}{2}-1}}{(mx+n)^{\frac{m+n}{2}}}2^{\frac{m+n}{2}}\Gamma\left(\frac{m+n}{2}\right)$$

将（3.5.9）式代入上式，即得到（3.5.8）式。

习　　题

1　已知随机变量在$[a\ b]$内均匀分布

（1）用几何方法说明$P(a\leqslant\xi\leqslant b)=1$

（2）求ξ落在$[a,c]$内的概率：

(a) 若$c<b$

(b) 若$c>b$

2　设ξ为连续型随机变量，其分布密度为

$$p(x)=\begin{cases}ax & 0\leqslant x\leqslant 1\\ a & 1<x\leqslant 2\\ -ax+3a & 2<x\leqslant 3\\ 0 & \text{其它}\end{cases}$$

（1）确定a值

（2）若 ξ_1,ξ_2,ξ_3是对ξ的三次独立观测值(观测值从理论上可看成与ξ同分布的随机变量）．求这三个观测值中正好一个大于1.5的概率．

3　设ξ是$[-a,a]$上均匀分布的随机变量，其中$a>0$,试确定满足下面关系的正数a

(1) $P(\xi>1)=\frac{1}{3}$　　(2) $P(\xi>1)=\frac{1}{2}$

(3) $P\left(\xi<\frac{1}{2}\right)=0.7$　　(4) $P\left(\xi<\frac{1}{2}\right)=0.3$

(5) $P(|\xi|<1)=P(|\xi|>1)$

4　设某电子管的寿命（小时）ξ是一个具有密度函数

$$f(x)=\begin{cases}100/x^2 & x>100\\ 0 & 其它\end{cases}$$

的连续型随机变量.

(1) 若已知这个电子管在工作了150小时以后仍然好使，求至少工作200小时的概率.

(2) 若有三个这种管子装在一个设备里，求在150小时之内恰有一个管子被换掉的概率.

(3) 问在一个设备里最多能装多少这种管子，能有$\frac{1}{2}$的概率保证它们都工作150小时以上.

5　设ξ的密度函数$p(x)$是原点对称的，$a>0$为一常数. 试用几何直观验证下面结果

(1) $F(0)=0.5$

(2) $P(\xi>a)=P(\xi\geqslant a)=0.5-\int_0^a p(x)dx$

(3) $F(-a)+F(a)=1$

(4) $P(-a<\xi<a)\equiv P(-a\leqslant\xi\leqslant a)=2F(a)-1$

(5) $P(|\xi|>a)\equiv P(|\xi|\geqslant a)=2F(-a)=2[1-F(a)]$.

6　连续型随机变量ξ的分布密度为$f(x)=3x^2$，$-1\leqslant x\leqslant 0$，对任一b：$-1<b<0$，计算 $P\left(\xi>b\,\middle|\,\xi<\frac{b}{2}\right)$

7　若f，g，均在同一区间$a\leqslant x\leqslant b$上是密度函数.

(1) 试证$f+g$不是这个区间上的密度函数.

(2) 对任一数 $\beta, 0<\beta<1$，$\beta f(x)+(1-\beta)g(x)$是这个区间上的密度函数

8　图（3—8题）中曲线表示一随机变量的分布密度，问a与

b 间有什么关系。

9　若连续随机变量 ξ 的分布函数与分布密度分别为 $F(x)$ 和 $f(x)$，求随机变量 $\eta=a\xi+b$ 的分布密度（其中 a，b 均为实数）。

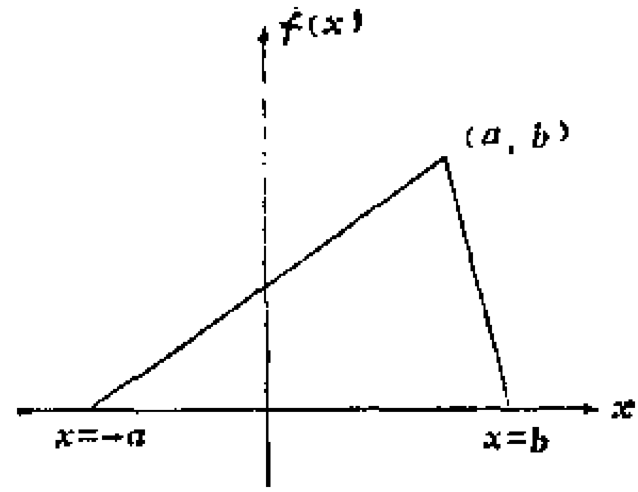


图　（3—8题）

10　设 ξ 在 $(-1,1)$ 上均匀分布，令 $\eta=4-\xi^2$，求 η 的分布密度并画出图来。

11　设 ξ 的分布密度为 $f(x)=e^{-x}, x>0$，求下列随机变量的分布密度

（1）$\eta=\xi^3$　　　　（2）$\zeta=3/(\xi+1)^2$

12　设随机变量 ξ 有连续分布函数 $F(x)=\int_{-\infty}^{x}p(t)dt$，试求下列随机变量的分布函数和密度

（1）$\eta=\xi^{-1}$　　　　（2）$\zeta=|\xi|$

13　设随机变数 ξ 具有单调上升的连续的分布函数 $F(x)$，问（i）$\eta=F(\xi)$ 是不是随机变量？为什么？（ii）如果是的话，它的分布函数是什么，它是什么分布随机变量。

14　若 $\xi\sim F(x)$，（密度为 $p(x)$）求 $\eta=\cos\xi$ 和 $\zeta=\mathrm{tg}\xi$ 的分布。

15　假设 ξ 是具有定义在 $(0,1)$ 上的某种分布的连续型随机变量，$P(\xi\leqslant0.29)=0.75$，又设 $\eta=1-\xi$，试确定使得 $P(\eta\leqslant k)=0.25$ 的 k 值。

16　设 ξ_1, ξ_2 分别表示某煤矿一月和二月份的产量，假如二维随机变量 (ξ_1,ξ_2) 的分布密度为

$$p(x,y)=\begin{cases} c & 5\leqslant x\leqslant10, 4\leqslant y\leqslant9 \\ 0 & \text{其他} \end{cases}$$

求 $\xi_1\geqslant\xi_2$ 的概率。

17　设二维随机变量 (ξ,η) 具有联合密度函数

$$f(x,y)=\begin{cases}kx(x-y) & 0<x<2 \quad -x<y<x \\ 0 & \text{其它}\end{cases}$$

（1）确定常数 k 值

（2）求 ξ,η 的边缘分布密度

18　设二维随机变量 (ξ,η) 的联合分布密度为

$$f(x,y)=\begin{cases}x^2+\dfrac{xy}{3} & 0<x<1 \quad 0<y<2 \\ 0 & \text{其　它}\end{cases}$$

计算（1）$P\left(\xi>\dfrac{1}{2}\right)$；　　（2）$P(\eta<\xi)$；

（3）$P\left(\eta<\dfrac{1}{2}\middle|\xi<\dfrac{1}{2}\right)$

19　设 $\xi_1,\xi_2,\cdots,\xi_n$ 是 n 个独立具有同样分布函数 $F(x)$ 的随机变量，$\xi_1^*\leqslant\xi_2^*\leqslant\cdots\leqslant\xi_k^*\leqslant\cdots\leqslant\xi_n^*$ 是它们的大小顺序，求 ξ_k^* 的分布函数。

20　随机变量 ξ,η 相互独立且分别在 $[0,1]$ 和 $[0,2]$ 上是均匀分布的，求 $\xi+\eta$ 的分布密度。

21　ξ 表示一电子设备的寿命，设 ξ 的分布密度为

$$f(x)=\begin{cases}\dfrac{1000}{x^2} & x>1000 \\ 0 & \text{其他}\end{cases}$$

又设 ξ_1,ξ_2 是对 ξ 的两次独立测定，求随机变量 $\eta=\dfrac{\xi_1}{\xi_2}$ 的分布密度。

22　设二维随机变量 (ξ_1,ξ_2) 的联合分布密度是

$$\frac{1}{\Gamma(k_1)\Gamma(k_2)}x_1^{k_1-1}(x_2-x_1)^{k_2-1}e^{-x_2} \quad 0<x_1<x_2<\infty$$

试求 ξ_1 和 ξ_2 的边缘密度。

23　在 $(0\ a)$ 线段上任意投掷两个点〔即其横坐标均匀分布于线段 $(0\ a)$ 内〕试求两点距离的分布函数。

24　当电流 I 通过电阻 R 时所做的功为 $W=I^2R$，假设 I 和 R

是相互独立的随机变量，其分布密度分别是

$$I: f(i)=\begin{cases}6i(1-i) & 0\leqslant i\leqslant 1\\ 0 & \text{其它}\end{cases}$$

$$R: g(r)=\begin{cases}2r & 0<r<1\\ 0 & \text{其它}\end{cases}$$

确定随机变量W的分布密度，并画出它的曲线。

25　电车每隔 3 分钟来一趟，求乘客的平均等车时间。

26　设ξ是一电子设备的寿命（以小时计），假设ξ的分布密度$f(x)=k/x^n$，$2000\leqslant x\leqslant 10000$

（1）对$n=2$　确定k。

（2）对$n=3$　确定k。

（3）对一般的n确定k。

（4）求上述三种情形下的均值。

27　假设电子设备的寿命ξ（以1000小时为单位）是一连续型随机变量，其分布密度为

$$f(x)=e^{-x}\quad x>0$$

又设制造一个这种元件的费用是2.00元，售价是5.00元，而且规定$\xi\leqslant 0.9$时保证退款，问制造者每个平均利润是多少？

28　设随机变量ξ，η相互独立，分布密度分别为

$$p_\xi(x)=\begin{cases}0 & x\leqslant 0\\ c_1x^\alpha e^{-\beta x} & x>0\end{cases}$$

$$p_\eta(x)=\begin{cases}0 & x\leqslant 0\\ c_2x^\gamma e^{-\beta x} & x>0\end{cases}$$

试求（1）常数c_1和c_2。

（2）$\xi+\eta$的分布密度。

（3）$\xi+\eta$的均值和方差

29　假设ξ和η相互独立，其分布密度分别为

$$f(x)=8/x^3\qquad x>2$$

$$g(y)=2y\qquad 0<y<1$$

（1）求$\zeta=\xi\cdot\eta$的分布密度

（2）用两种方法计算$E\zeta$：（i）利用（1）中得到的密度函数；(ii) 不用ζ的密度函数而直接计算．

30 设随机变量ξ服从对数正态分布，即ξ的分布密度为

$$p(x)=\frac{1}{x\beta\sqrt{2\pi}}e^{-\frac{1}{2\beta^2}(\ln x-\alpha)^2}\quad x>0$$

求$E\xi$和$D\xi$．

31 设$\xi_1,\xi_2,\cdots,\xi_n$相互独立同均匀分布(如在$(0,\alpha)$区间上)，令$\xi_1^*=\min\limits_{1\leqslant i\leqslant n}\{\xi_i\}$，$\xi_n^*=\max\limits_{1\leqslant i\leqslant n}\{\xi_i\}$求$\xi_1^*$和$\xi_n^*$的均值和方差．

32 设二维随机变量（ξ，η）在图（3—32题）的三角区域内均匀分布，求 $D\xi$和 $D\eta$.

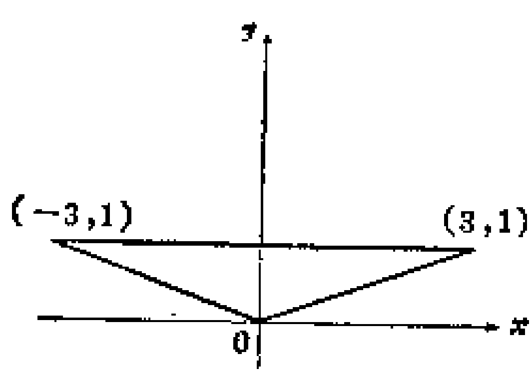


图 （3—32题）

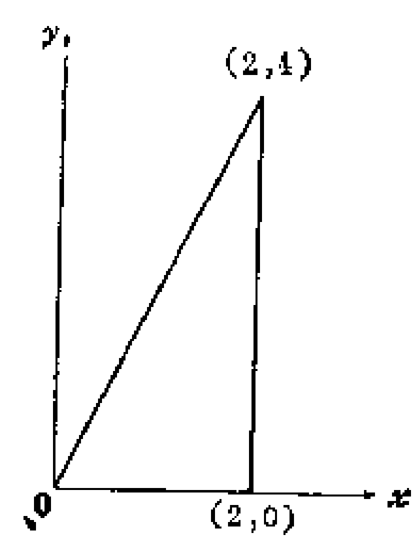


图 （3—33题）

33 假设（ξ，η）在图(3—33题)的三角区域内均匀分布

（1）求ξ和η的边缘密度

（2）计算出$D\xi$和$D\eta$

34 若$\xi\sim N(2,0.16^2)$，利用正态表估计下列概率：

（1）$P\{\xi\geqslant 2.3\}$

（2）$P\{1.8\leqslant\xi\leqslant 2.1\}$

35 检查测高器的精确性，结果表明误差不超过±2.5米的情况占所有情况的90%，设测高器的误差服从N（$0,\sigma^2$），求σ．

36　电缆的直径服从均值为0.8方差为0.0004的正态分布，求直径超过0.81的概率。

37　我们可能感兴趣的只是ξ的大小，即$\eta=|\xi|$，若$\xi\sim N(0,1)$，试确定η的分布密度并计算$E\eta$和$D\eta$。

38　设$\xi\sim N(\mu,\sigma^2)$，确定c（作为μ和σ的函数）使得$P(\xi\leqslant c)=2P(\xi>c)$。

39　设随机变量ξ的密度为

$$p(x)=\frac{1}{2a}exp\left\{-\frac{|x-a|}{a}\right\}$$

（拉普拉斯分布）试求$E\xi$和$D\xi$。

40　设$\xi\sim N(0,5^2)$。估计$P(1<\xi^2<4)$。

41　设一块厚度为T的金属板上的微观坑的深度为ξ，假设ξ服从参数为α的指数分布，求$E\xi$。

42　设ξ服从$N(0,1)$，ξ_1和ξ_2是对ξ的两次独立观测，求$Z=\xi_1+\xi_2, W=\xi_1-\xi_2$的联合分布并证$Z$与$W$相互独立。

43　设ξ_1,ξ_2同题42，求样本平均$\overline{\xi}=(\xi_1+\xi_2)/2$与离差平方和$S^2=(\xi_1-\overline{\xi})^2+(\xi_2-\overline{\xi})^2$的联合及边缘分布密度，验证$\overline{\xi}$与$S^2$的独立性，并求$\zeta=\overline{\xi}\sqrt{2}/\sqrt{S^2}$的分布。

学习指导

（一）本章概要

连续型随机变量不管从理论方面还是从应用角度来看都是一种非常重要的随机变量类型，所以必须掌握描述连续型随机变量取值规律的分布密度和分布函数以及其均值、方差等数字特征等概念的定义实质。

最常见最基本的连续型随机变量有均匀分布、指数分布和正态分布随机变量，这些随机变量的分布密度、分布函数以及其中所包

含的参数的涵义必须熟记清楚。

要掌握二维连续型随机变量的定义，注意联合分布与边缘分布的关系及独立性等概念。

随机变量的函数的学习要掌握如何从已知分布求出做为函数的随机变量的分布的方法。对于在理论上及在数理统计里有着广泛应用的三个重要的随机变量的函数χ^2，t，F分布，应掌握住如何去判断一个随机变量的函数是否服从这些分布。

（二）内容说明

1 要正确理解分布密度的概念

为什么在离散型随机变量时，刻划其概率规律使用的是分布列，而在连续型随机变量场合要用分布密度来刻划呢？要想正确理解分布密度的概念，必须仔细思考这个问题。

不同性质的事物用不同的方法去研究，用不同的工具去解决。举个通俗的例子来说，在市场上，香瓜可以按个出售，也可以论斤卖，而猪肉只能论斤卖了。这是因为香瓜的重量是按个分布的，只要我们知道了每个香瓜的重量，则任意几个香瓜的重量就会知道了。可是猪肉就没有象香瓜这样“个”的单位，不可能同样对待。如果我们把离散型随机变量想象成类似香瓜这类东西，而连续型随机变量比成象猪肉这类物品，概率分布相应地理解成“重量”分布，则概率分布密度就相当于物体的质量线密度函数。

密度函数在整个数轴积分等于1这一性质，在一些练习题里常常被用来确定密度函数中的未知常数。比如，设随机变量的密度函数形如

$$p(x)=\begin{cases} x & 0\leqslant x\leqslant 1 \\ a-x & 1<x\leqslant 2 \\ 0 & \text{其它} \end{cases} \tag{3—1}$$

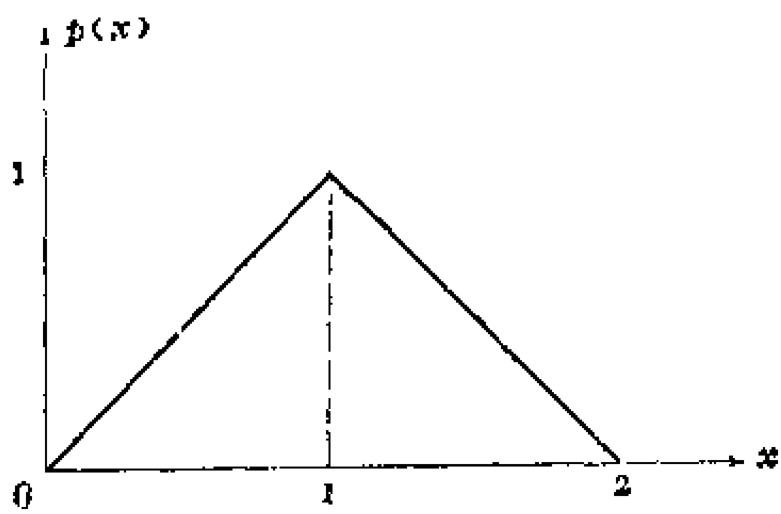
试确定a值。由于

$$1=\int_{-\infty}^{\infty}p(x)dx=\int_0^1 xdx+\int_1^2(a-x)dx$$

$$=\frac{x^2}{2}\Big|_0^1+\left(ax-\frac{x^2}{2}\right)\Big|_1^2$$

$$=\frac{1}{2}+(2a-2)-\left(a-\frac{1}{2}\right)=a-1$$

故$a=2$。（3—1）式确定的实际是三角形分布．见图（3—1）。



图（3—1）

对于连续型随机变量，我们要牢记：随机变量取值于某个区间的概率就等于密度函数在此区间上的积分．比如在上例中，我们要计算事件$(1.5<\xi\leqslant 2)$的概率，则

$$P(1.5<\xi\leqslant 2)=\int_{1.5}^{2}p(x)dx=\int_{1.5}^{2}(2-x)dx=0.125$$

2　常用连续型随机变量综合列表（见第212页）

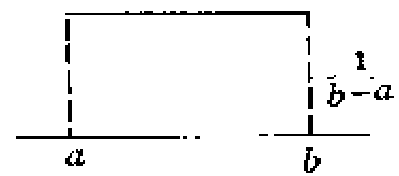
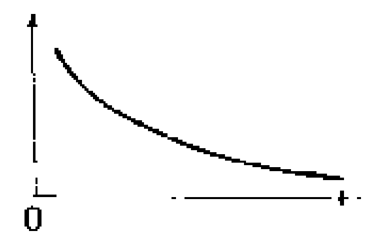
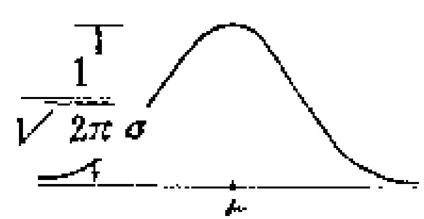
3　随机变量函数的分布的计算

通过ξ的分布计算函数$\eta=g(\xi)$的诱导分布的问题，当相应的定理的条件满足时，可以利用定理计算．但在一般情况下往往是设法直接计算η的分布函数

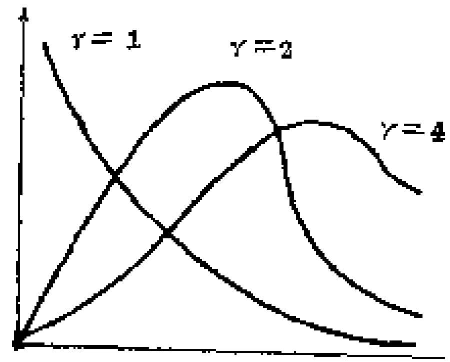
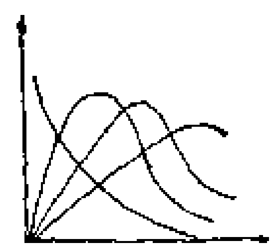
$$G(y)=P(\eta\leqslant y)=P(g(\xi)\leqslant y)$$
$$=P(\xi\in g^{-1}(-\infty,y))\qquad（3—2）$$

其中g^{-1}为g的逆变换．只不过是当定理3.3.1的条件满足时，事件

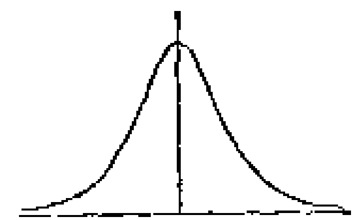
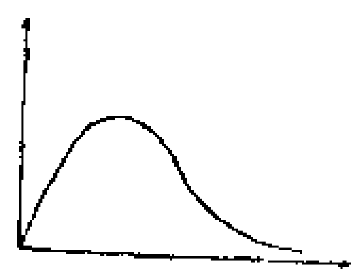
常用连续型随机变量综合列表

分布名称	密度函数 $f(x)$	$f(x)$的图形	矩母函数*	均值方差	附注
均匀分布	$f(x)=\begin{cases}\frac{1}{b-a} & a<x<b\\ 0 & \text{其他}\end{cases}$		$\frac{1}{(b-a)t}(e^{bt}-e^{ta})$	$E\xi=\frac{a+b}{2}$ $D\xi=\frac{(b-a)^2}{12}$	
指数分布	$f(x)=\begin{cases}ae^{-ax} & x\geqslant 0\\ 0 & \text{其他}\end{cases}$		$\frac{a}{a-t}$	$E\xi=\frac{1}{a}$ $D\xi=\frac{1}{a^2}$	指数分布是Gamma分布的特殊情形
正态分布 $N(\mu,\sigma^2)$	$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ $-\infty<x<\infty$		$e^{t\mu+\sigma^2t^2/2}$	$E\xi=\mu$ $D\xi=\sigma^2$	加法定理成立 如ξ_i独立服从$N(\mu_i,\sigma_i^2)$，c_i，d为常数，则 $\sum_{i=1}^{n}c_i\xi_i+d\sim N\left(\sum_{i=1}^{n}c_i\mu_i+d,\ \sum_{i=1}^{n}c_i^2\sigma_i^2\right)$

*第四章讲

Gamma 分布	$f(x) = \begin{cases} \frac{\alpha}{\Gamma(\gamma)}(\alpha x)^{\gamma-1}e^{-\alpha x} & 0<x<\infty \\ 0 & 其他 \end{cases}$		$\left(\frac{\alpha}{\alpha-t}\right)^{\gamma}$	$E\xi=\frac{\gamma}{\alpha}$ $D\xi=\frac{\gamma}{\alpha^2}$	1. $\gamma=1$时，Gamma分布为指数分布。 2. 加法定理成立：若ξ, η独立，$\xi\sim G(\alpha,\gamma_1)$, $\eta\sim G(\alpha,\gamma_2)$则$\xi+\eta\sim G(\alpha,\gamma_1+\gamma_2)$ 3. 当$\alpha=\frac{1}{2}$, $\gamma=\frac{n}{2}$, n为正整数，化为$\chi^2(n)$分布
$\chi^2(n)$分布	$f(x) = \begin{cases} \frac{1}{\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}}}x^{\frac{n}{2}-1}e^{-\frac{x}{2}} & x>0 \\ 0 & 其他 \end{cases}$ n为正整数		$\left(\frac{\frac{1}{2}}{\frac{1}{2}-t}\right)^{\frac{n}{2}}$	$E\xi=n$ $D\xi=2n$	1. 加法定理成立，独立的χ^2分布随机变量之和还是χ^2分布，其自由度为加项自由度之和 2. 如果ξ_i独立同N(0,1)，则$\sum_{i=1}^{n}\xi_i^2\sim\chi^2(n)$

续表

分布名称	密度函数 $f(x)$	$f(x)$的图形	矩母函数*	均值方差	附　　注
t 分布	$f(t)=\dfrac{\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\left(1+\dfrac{t^2}{2}\right)^{-\frac{n+1}{2}}$ $-\infty<t<\infty$ n为正整数			$E\xi=0$ $(n>1)$ $D\xi=\dfrac{n}{n-2}$ $(n>2)$	设 $\xi,\xi_1,\cdots,\xi_n$ 独立同 $N(0,\sigma^2)$，则 $\xi/\sqrt{\dfrac{1}{n}\sum\limits_{i=1}^{n}\xi_i^2}\sim t(n)$
F分布 $F(m,n)$	$f_F(x)=\begin{cases}\dfrac{\left(\frac{m}{n}\right)^{\frac{m}{2}}x^{\frac{m}{2}-1}}{B\left(\frac{m}{2},\frac{n}{2}\right)}\left(1+\dfrac{m}{n}x\right)^{-\frac{m+n}{2}} & x>0\\ 0 & \text{其他}\end{cases}$	其他 			如ξ,η独立 $\xi\sim\chi^2(m)$ $\eta\sim\chi^2(n)$ 则 $\dfrac{\xi/m}{\eta/n}\sim F(m,n)$

$(\xi\in g^{-1}(-\infty,y))$根据$g$是递增或递减变成事件$(\xi\leqslant g^{-1}(y))$或$(\xi\geqslant g^{-1}(y))$而已。所以不管定理条件是否满足，我们都可通过(3—2)式去计算η的分布函数。

根据二维随机变量(ξ,η)的分布计算二元随机变量函数的诱导分布的问题，可以直接先去计算函数的分布函数，然后求分布密度，也可以利用定理3.3.3去计算。利用此定理的关键是要根据不同的问题，设出另一个二元函数。这样先计算作为函数的二维随机变量的联合分布密度，再求边缘分布就可得出所需的结果。如果所设的二元函数合适，就会使问题计算变得简单。

例3.1　若$\xi\sim\chi^2(m)$，$\eta\sim\chi^2(n)$，ξ，η相互独立，求$F=\dfrac{\xi/m}{\eta/n}$的分布。我们已经证明了$F\sim F(m,n)$，现在直接利用定理3.3.3证明。设变换

$$u=x+y,\ v=\frac{x}{y}\frac{n}{m}, \text{则} u\geqslant 0,\quad v\geqslant 0,\ \text{且}$$

$$x=\frac{uv\dfrac{m}{n}}{1+\dfrac{m}{n}v},\quad y=\frac{u}{1+\dfrac{m}{n}v}$$

$$J^{-1}=\frac{\partial(u,v)}{\partial(x,y)}=\begin{vmatrix}\dfrac{\partial u}{\partial x} & \dfrac{\partial u}{\partial y}\\ \dfrac{\partial v}{\partial x} & \dfrac{\partial v}{\partial y}\end{vmatrix}=\begin{vmatrix}1 & 1\\ \dfrac{1}{y}\dfrac{n}{m} & -\dfrac{x}{y^2}\dfrac{n}{m}\end{vmatrix}$$

$$=-\frac{x+y}{y^2}\frac{n}{m}=-\frac{\left(1+\dfrac{m}{n}v\right)^2}{u}\frac{n}{m}$$

所以　$$|J|=\frac{m}{n}\frac{u}{\left(1+\dfrac{m}{n}v\right)^2}$$

令　$\zeta=\xi+\eta$，$F=\dfrac{\xi/m}{\eta/n}$，则(ζ,F)的联合分布密度

$$g(u,v)=\frac{1}{\Gamma\left(\frac{m}{2}\right)2^{\frac{m}{2}}}x^{\frac{m}{2}-1}e^{-\frac{x}{2}}\cdot\frac{1}{\Gamma\left(\frac{n}{2}\right)2^{\frac{n}{2}}}y^{\frac{n}{2}-1}e^{-\frac{y}{2}}$$

$$\cdot\frac{m}{n}\frac{u}{\left(1+\frac{m}{n}v\right)^2}$$

$$=\frac{e^{-\frac{x+y}{2}}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)2^{\frac{m+n}{2}}}\left(\frac{uv\frac{m}{n}}{1+\frac{m}{n}v}\right)^{\frac{m}{2}-1}\left(\frac{u}{1+\frac{m}{n}v}\right)^{\frac{n}{2}-1}$$

$$\cdot\frac{m}{n}\frac{u}{\left(1+\frac{m}{n}v\right)^2}$$

$$=\frac{\left(\frac{m}{n}\right)^{\frac{m}{2}}v^{\frac{m}{2}-1}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{m}{n}v\right)^{-\frac{m+n}{2}}u^{\frac{m+n}{2}-1}e^{-\frac{u}{2}}$$

所以F的边缘分布密度

$$g(v)=\frac{\left(\frac{m}{n}\right)^{\frac{m}{2}}v^{\frac{m}{2}-1}}{\Gamma\left(\frac{m}{2}\right)\Gamma\left(\frac{n}{2}\right)}\left(1+\frac{m}{n}v\right)^{-\frac{m-n}{2}}\int_0^\infty u^{\frac{m+n}{2}-1}e^{-\frac{u}{2}}du$$

$$=\frac{\left(\frac{m}{n}\right)^{\frac{m}{2}}v^{\frac{m}{2}-1}}{B\left(\frac{m}{2},\frac{n}{2}\right)}\left(1+\frac{m}{n}v\right)^{-\frac{m+n}{2}}$$

利用定理3.3.4计算两个随机变量ξ，η的和、差、积、商的分布时，对于那种ξ、η的密度函数是分段给出的场合，读者要特别注意，这里的关键是要首先设法确定出随机变量函数的分布密度表达式里被积函数取值的非零区域。这一步对初学者往往是难点。

例3.2　若ξ，η相互独立，分布密度分别为

$$p_1(x)=\begin{cases}\frac{1}{2} & 0\leqslant x\leqslant 3\\ 0 & \text{其他}\end{cases}$$

$$p_2(x)=\begin{cases}e^{-(x-2)} & x>2\\ 0\ , & \text{其他}\end{cases}$$

求 $W=\xi/\eta$ 的分布。

由定理3.3.4，W的分布密度

$$p_W(w)=\int_{-\infty}^{\infty}p_1(xw)p_2(x)|x|dx \qquad (3-3)$$

当且仅当$p_1(xw)\neq 0$同时 $p_2(x)\neq 0$，即满足不等式$1\leqslant xw\leqslant 3$，$x>2$时被积函数非零。见图（3－2）中的阴影部分。

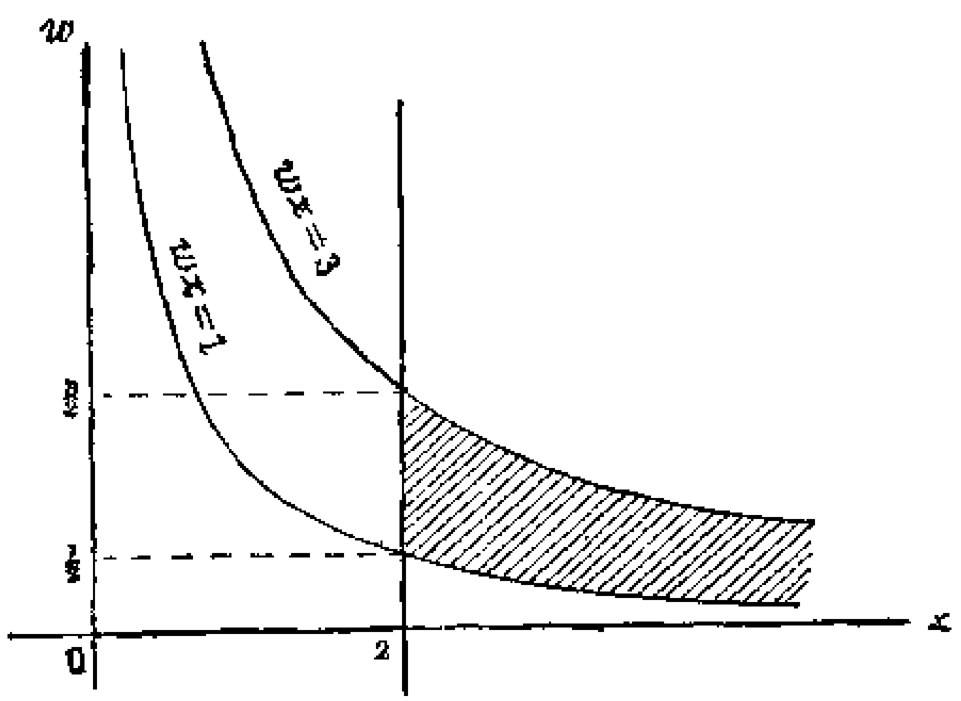


图 （3－2）

由图可以看出，对于大于$\frac{3}{2}$和小于0的w，（3－3）式积分中的被积函数恒为0，所以这时$p_W(w)=0$。当$\frac{1}{2}\leqslant w\leqslant\frac{3}{2}$时，

$$p_W(w)=\int_2^{\frac{3}{w}}\frac{1}{2}xe^{-(x-2)}dx=\frac{-e^2}{2}\left[e^{-x}(x+1)\right]_2^{\frac{3}{w}}$$

$$=\frac{3}{2}-\frac{e^2}{2w}e^{-3/w}(3+w)$$

当 $0<w<\frac{1}{2}$ 时，

$$p_W(w)=\int_{\frac{1}{w}}^{\frac{3}{w}}\frac{1}{2}xe^{-(x-2)}dx=\frac{-e^2}{2}[e^{-x}(x+1)]_{\frac{1}{w}}^{\frac{3}{w}}$$

$$=\frac{e^2}{2w}[e^{-\frac{1}{w}}(1+w)-e^{-\frac{3}{w}}(3+w)]$$

4 应用问题举例

例3.3 设已知元件的寿命T服从指数分布，而生产一种元件的费用c是元件平均寿命$E(T)=\mu$的函数

$$c=3\mu^2$$

设元件在使用过程中每小时能创造D元价值，这样，能使每个元件为国家积累 $p=DT-3\mu^2$元财富，其中T是元件的寿命.

$$E(p)=D\mu-3\mu^2$$

我们的问题是，生产平均寿命是多少的元件为国家贡献最大呢？这就要寻求使得$E(p)$达到极大的μ值。令

$$\frac{dE(p)}{d\mu}=0$$

解得$\mu=D/6$。这时可使每个元件的贡献为

$$E(p)_{max}=D^2/12$$

例3.4 已知保险丝的寿命服从指数分布，制造它有两种工艺：工艺Ⅰ生产的产品平均寿命为100小时（参数为100^{-1}），工艺Ⅱ为150小时（参数为150^{-1}）．工艺Ⅰ每个产品造价C元，工艺Ⅱ每个产品造价为工艺Ⅰ的2倍．如果保险丝不能使用200小时以上，还要支付K元损失，问应采用那种工艺.

我们先计算每个工艺产品的平均成本，对工艺Ⅰ，设成本为C_1，则

$$C_1=\begin{cases}C & 若T>200\\ C+K & 若T\leqslant 200\end{cases}$$

因此 $$EC_{\text{I}}=C\cdot P(T>200)+(C+K)P(T\leqslant 200)$$

$$=Ce^{-(\frac{1}{100})200}+(C+K)(1-e^{-(\frac{1}{100})200})$$

$$= Ce^{-2} + (C+K)(1-e^{-2}) = K(1-e^{-2}) + C$$

同样可算得

$$EC_{\mathrm{I}} = K(1-e^{-\frac{4}{3}}) + 2C$$

这样 $\quad EC_{\mathrm{I}} - EC_{\mathrm{I}} = C + K(e^{-2} - e^{-\frac{4}{3}}) \doteq C - 0.13K$

所以倘若$C>0.13K$，我们就选取工艺Ⅰ.

例3.5　电子管的寿命T服从参数为β的指数分布. 机器使用这种电子管每小时花c_1元电费，创造c_2元价值，假如每隔H小时换一支新管子，规定每小时折旧费为c_3元，问更换时间H应多长才最合算.

用R表示获益多少，则

$$R = \begin{cases} c_2H - c_1H - c_3H & \text{若} T>H \\ c_2T - c_1T - c_3H & \text{若} T\leqslant H \end{cases}$$

R是寿命T的函数，记为$R(T)$

$$ER(T) = \int_0^{\infty} R(t)\beta e^{-\beta t}dt = \int_H^{\infty} R(t)\beta e^{-\beta t}dt + \int_0^{H} R(t)\beta e^{-\beta t}dt$$

所以

$$\begin{aligned} E(R) &= H(c_2-c_1-c_3)P(T>H) - c_3HP(T\leqslant H) \\ &\quad + (c_2-c_1)\int_0^H t\beta e^{-\beta t}dt \\ &= H(c_2-c_1-c_3)e^{-\beta H} - c_3H(1-e^{-\beta H}) \\ &\quad + (c_2-c_1)[\beta^{-1} - e^{-\beta H}(\beta^{-1}+H)] \\ &= (c_2-c_1)[He^{-\beta H} + \beta^{-1} - e^{-\beta H}(\beta^{-1}+H)] - c_3H \end{aligned}$$

为求$E(R)$极大值，解下列方程：

$$\begin{aligned} \frac{dE(R)}{dH} &= [c_2-c_1][H(-\beta)e^{-\beta H} + e^{-\beta H} - e^{-\beta H} \\ &\quad + (\beta^{-1}+H)\beta e^{-\beta H}] - c_3 \\ &= (c_2-c_1)e^{-\beta H} - c_3 = 0 \end{aligned}$$

得

$$H = -\left(\frac{1}{\beta}\right)\ln\left[\frac{c_3}{c_2-c_1}\right]$$

为使此解有意义必须有$H>0$，即$0<c_3/(c_2-c_1)<1$，这等价于$c_2-c_1>c_3$而$c_2-c_1-c_3>0$。

特别设 $\beta=0.01$，$c_1=3$元 $c_2=10$元 $c_3=4$元，则 $H=-100\ln\left(\frac{4}{7}\right)=55.9\doteq56$。这说明每隔56小时更换一次最有利。

*例3.6 考虑一个厚度为 s 的物体，其表面有不同深度的微观坑，而这些坑的深度由于化学腐蚀又在增加，每当有一个坑的深度穿透了物体就发生了损坏。假设开始的坑深服从指数分布，并假定穿透时间和物体去掉此坑后剩余的厚度成正比，确定损坏时间的分布。

设有坑N个，用t_i表示第 i 个坑的初始深度，则$\tau_i=K(s-t_i)$为第 i 个坑的穿透时间，其中$K>0$

由于t_i不大于 s，故 t_i的分布密度为具有参数 λ 的指数截尾分布

$$f(t)=\begin{cases}0 & t>s\\ c\lambda e^{-\lambda t} & 0\leqslant t<s\end{cases}$$

$$1=\int_0^s f(t)dt=\int_0^s c\lambda e^{-\lambda t}dt=-c(e^{-\lambda s}-1)$$

$$\therefore\quad c=\frac{1}{1-e^{-\lambda s}}$$

分布函数为

$$F(t)=\frac{1}{1-e^{-\lambda s}}(1-e^{-\lambda t})$$

$$P(t_i>t)=\frac{e^{-\lambda t}-e^{-\lambda s}}{1-e^{-\lambda s}}\qquad 0\leqslant t\leqslant s$$

现在令$G(\tau)=P(\tau_i\leqslant\tau)=P\left(t_i\geqslant s-\frac{\tau}{K}\right)=\frac{e^{\lambda\tau/k}-1}{e^{\lambda s}-1}\quad 0\leqslant\tau\leqslant Ks$。设$T$为物体损坏时间，则

$$T=\min_{1\leqslant i\leqslant N}(\tau_i)$$

---

* 例3.6不是必读内容。

$$F_T(x)=P(T\leqslant x)=1-[1-G(x)]^N$$

*例3.7　考虑由于疲劳裂纹而产生损坏的物理过程。

用随机序列$x_1<x_2<\cdots<x_n$表示物体在相邻阶段上生长的疲劳裂纹的大小。假定裂纹达到x_n，元件便发生损坏。又设第 i 步裂纹的增长x_i-x_{i-1}随机正比于裂纹x_{i-1}的大小，即

$$x_i-x_{i-1}=\pi_i x_{i-1}$$

其中 $\pi_i(i=1,2,\cdots,n)$为独立同分布随机变量。x_0 可解释为元件里隐含的微小裂缝或空隙。这样

$$\frac{x_i-x_{i-1}}{x_{i-1}}=\pi_i$$

$$\sum_{i=1}^{n}\pi_i=\sum_{i=1}^{n}\frac{x_i-x_{i-1}}{x_{i-1}}$$

假定$x_i-x_{i-1}=\Delta x_{i-1}$对每一步 i 都是很小的，

$$\sum_{i=1}^{n}\pi_i=\sum_{i=1}^{n}\frac{\Delta x_i}{x_{i-1}}\xrightarrow[\left(\substack{\Delta x\to 0\\ n\text{ 变大}}\right)]{}\int_{x_0}^{x_n}\frac{1}{x}dx=\ln x_n-\ln x_0$$

或
$$\ln x_n=\sum_{i=1}^{n}\pi_i+\ln x_0$$

由中心极限定理知(见§4.3)　$\sum\limits_{i=1}^{n}\pi_i$渐近正态，因此$x_n$的密度函数

$$g(x)=\begin{cases}\dfrac{1}{\sqrt{2\pi}\,\sigma x}exp\left[-\dfrac{1}{2}\left(\dfrac{\ln x-\ln x_0-\mu}{\sigma}\right)^2\right] & x>0\\ 0 & \text{其它}\end{cases}$$

渐近称为对数正态分布。

5　随机变量间的几种关系和分类

我们学过了很多种重要的连续型随机变量，比如，均匀分布、指数分布、正态分布、加马分布、χ^2分布、t－分布和 F－分布等

---

* 例3.7不是必读内容

等．这些随机变量之间有着纵横交错的联系和关系．为了能够很好的掌握这些随机变量，有必要从不同的角度把它们分类，来总结和概括它们之间的这些联系．

1　基本的和非基本的随机变量．在连续型随机变量中，有些随机变量就其理论地位来说是相对基本的，起着基础的作用，我们姑且把它们称之为基本随机变量．而其他的则或者是基本随机变量的特殊情形，或者是它们的某种函数，所以后一种我们称之为非基本的随机变量．显然，在我们所学的连续型随机变量中，均匀分布、正态分布、加马分布随机变量是基本随机变量，其他的则是非基本的．比如指数分布、χ^2分布乃是加马分布的特殊情形．$t-$分布，χ^2分布又是独立的正态分布随机变量的函数．$F-$分布是独立的χ^2分布的函数，所以也可以看成是正态分布随机变量的函数．

2　均匀分布和非均匀分布随机变量．如果从随机变量的分布是否是均匀的（即随机变量取值于某一区间的概率与此区间的长度成正比）来划分连续型随机变量的话，我们可以把均匀分布随机变量独分一类，其他的为另一类．这是一个看起来不十分协调的分类方法，一类似乎太小，而另一类又似乎太大．可是这一小一大之间却存在着下面有趣的联系：

若ξ是具有分布函数为$F(x)$的随机变量，则$F(\xi)$便为$[0,1]$上均匀分布随机变量．反之，若F为任一分布函数，η为$[0,1]$上均匀分布随机变量，则一定存在一个函数h，使得$h(\eta)$为具有分布函数F的随机变量，即

$$P(h(\xi)\leqslant x)=F(x),\quad x\in(-\infty,\infty)$$

证明　设$F(x)$为ξ的分布函数，因为$F(x)$是非降函数，对任意$0\leqslant y\leqslant 1$，可定义

$$F^{-1}(y)=\inf\{x:E(x)\geqslant y\}$$

我们有$\{F^{-1}(y)\leqslant x\}=\{y\leqslant F(x)\}$．这是因为由$F^{-1}(y)$的定义，从$y\leqslant F(x)$可立刻推出$F^{-1}(y)\leqslant x$．反之，$F^{-1}(y)\leqslant x$意味着对每个$\varepsilon>0$，有$y\leqslant F(x+\varepsilon)$，由于$\varepsilon$是任意的且$F$是右连续的，所以当

$\varepsilon \to 0$时，有 $y \leqslant F(x)$。

下面考查随机变量$\zeta = F(\xi)$的分布。对$0 \leqslant x \leqslant 1$，

$$P(\zeta \leqslant x) = P(F(\xi) \leqslant x)$$
$$= F(\xi \leqslant F^{-1}(x)) = F(F^{-1}(x)) = x$$

即$\zeta = F(\xi)$服从〔0,1〕均匀分布。

反之，设F为任一分布函数，η在〔0,1〕上均匀分布。令$h(x) = F^{-1}(x)$，考虑随机变量$h(\eta)$。

$$P(h(\eta) \leqslant x) = P(F^{-1}(\eta) \leqslant x)$$
$$= P(\eta \leqslant F(x)) = F(x)$$

上面给出的均匀分布与其他分布之间的关系在实际中有着重要的应用，因为根据这一关系，只要能设法产生出〔0,1〕上均匀分布的随机样本（均分分布随机数），就可产生出任意分布的随机样本（随机数）。这个结论在'蒙特卡洛'方法中是非常基本的。

3　加法定理成立的随机变量。两个同类型分布的独立随机变量，如果其和随机变量的分布不改变分布类型，则称对这种随机变量加法定理成立。对加马分布(包括χ^2分布)、正态分布加法定理都是成立的。均匀分布对加法定理显然不满足，因为两个均匀分布随机变量之和不再是均匀分布而变成了三角形分布。有趣的是，从第四章将要学习的中心极限定理可知，不管是对加法定理成立的随机变量，还是不成立的随机变量，独立同分布随机变量之和都以正态分布做为极限分布。也就是说，尽管对那些满足加法定理的随机变量，其和的分布还保持本来的分布类型，但只要加项无限增加时，它们都与正态分布相差不多了。比如独立同加马分布随机变量和尽管还服从加马分布，但只要求和项增大，这时的加马分布就渐近于正态分布了。

6　正态分布、指数分布和加马分布的识别

正态分布、指数分布和加马分布都是在实际中有着广泛应用的随机变量，所以学会在实际问题中识别它们是很重要的。

1　正态分布的识别方法。正态分布是概率论中最重要的一种

分布，它在自然界中是最常见的。如测量误差，零件的尺寸、重量，炮弹弹着点的坐标，某月的平均气温、雨量等等都可近似看成服从正态分布。识别一个随机变量是否服从正态分布，可采用定量的和定性的两种不同的方法。所谓定量的方法，就是首先要对随机变量进行观测，然后根据观测数据画出直方图，如果其直方图的形状为"中间高，两边低，呈对称"的样子，我们就可大致断言它是服从正态分布的。当然这是一种粗略的分析方法，要想根据观测数据进行科学分析，就需使用第五章讲述的统计检验的方法。

什么是识别正态分布的定性的方法呢？就是要首先对随机变量进行定性的分析。如果分析的结果可以认定这个随机变量是由大量的随机因素所影响，而每一个因素对总体所起作用又不太大，则可以断言这个随机变量就是服从正态分布的。这一点可以利用第四章讲的中心极限定理来说明。

2 指数分布的识别方法。指数分布是在可靠性理论里最常见的一种寿命分布。那么什么样的元件寿命服从指数分布呢？为了说明这个问题，我们先来考虑一般元件的所谓瞬时损坏的问题，即在某个确定时刻损坏的问题。设元件的寿命为 T，显然 T 为一连续型随机变量。所谓元件在时刻 t 损坏，就是指在 t 以前它还好使，而在其后的一瞬间失效。所以在 t 时刻损坏的概率指的就是下面的条件概率，

$$
\begin{aligned}
&P(t<T\leqslant t+\Delta t\mid T>t)\\
&=\frac{P(t<T\leqslant t+\Delta t)}{P(T>t)}=\frac{\int_t^{t+\Delta t}f(x)\,dx}{P(T>t)}\\
&=f(t')\Delta t/R(t)
\end{aligned}
$$

其中 $t'\in[t,t+\Delta t]$，$R(t)=P(T>t)=1-F(t)$，而 $f(x)$ 和 $F(x)$ 分别为 T 的密度函数和分布函数。令

$$
h(t)=\frac{f(t)}{R(t)}
$$

则

$$P(t<T\leqslant t+\Delta t|T>t)\doteq h(t)\Delta t$$

称函数$h(t)$为此元件的瞬时损坏率。容易证明

$$f(t)=h(t)exp\left\{-\int_0^t h(s)ds\right\} \qquad (3-4)$$

事实上，由于$R'(t)=-f(t)$，则$h(t)=-\dfrac{R'(t)}{R(t)}$把此式两边从0到$t$进行积分，得

$$\int_0^t h(s)ds=-\int_0^t \frac{R'(s)}{R(s)}ds=-\ln R(s)\Big|_0^t$$
$$=-\ln R(t)+\ln R(0)=-\ln R(t)$$

故有 $$R(t)=exp\left\{-\int_0^t h(s)ds\right\}$$

于是 $$f(t)=h(t)R(t)=h(t)exp\left\{-\int_0^t h(s)ds\right\}$$

现在我们来考查诸如保险丝、宝石轴承、玻璃制品这类元件，看看它们的使用寿命服从什么分布。容易发现，这类元件寿命的共同特点就是如果发现在t时刻元件没有损坏，那么在t以后继续使用时就象使用新的一样。也就是说条件概率$P(t<T\leqslant t+\Delta t|T>t)$应与$t$无关，即$h(t)\Delta t$应与$t$无关，或说$h(t)$应为常数$\lambda$。元件的寿命若具有此性质，我们称之为是具有‘无记忆性’的。由（3－4）式，此种元件的寿命T的密度函数

$$f(t)=\lambda e^{-\lambda t},\ t>0。$$

这正是具有参数为λ的指数分布。因此得出结论，若元件的寿命是具有无记忆性的，或说其瞬时损坏率$h(t)$为常数时，则它就服从指数分布。

3　加马分布的识别。我们还是讨论元件的寿命问题。设元件是在受到k次“冲击”或“骚动”时才发生损坏，而“冲击”发生的数目是服从参数为λ的普阿松分布，那么这时元件的寿命$T^{(k)}$服从什么分布呢？为了求$T^{(k)}$的分布密度$f_{T^{(k)}}(t)$，考虑概率

$$P(t<T^{(k)}\leqslant t+\Delta t)$$
$=P$(在$[0,t]$里恰发生$k-1$次冲击，而在$[t,t+\Delta t]$里发生

1次冲击)

$$=\frac{e^{-\lambda t}(\lambda t)^{k-1}}{(k-1)!}\cdot\lambda\Delta te^{-\lambda\Delta t}=\frac{e^{-\lambda t}\lambda^k t^{k-1}}{(k-1)!}\Delta te^{-\lambda\Delta t}$$

所以

$$f_{T^{(k)}}(t)=\lim_{\Delta t\to 0}\frac{P(t<T^{(k)}\leqslant t+\Delta t)}{\Delta t}$$

$$=\frac{e^{-\lambda t}\lambda^k t^{k-1}}{(k-1)!}\lim_{\Delta t\to 0}e^{-\lambda\Delta t}$$

$$=\frac{e^{-\lambda t}\lambda^k t^{k-1}}{\Gamma(k)}$$

这恰是参数为 λ 和 k 的加马分布。

习 题 解 答

1　(1) ξ 在$[a,b]$上均匀分布，则其分布密度为

$$p(x)=\begin{cases}\dfrac{1}{b-a} & a\leqslant x\leqslant b\\ 0 & \text{其他}\end{cases}$$

$p(x)$的曲线如图(3—1题)。

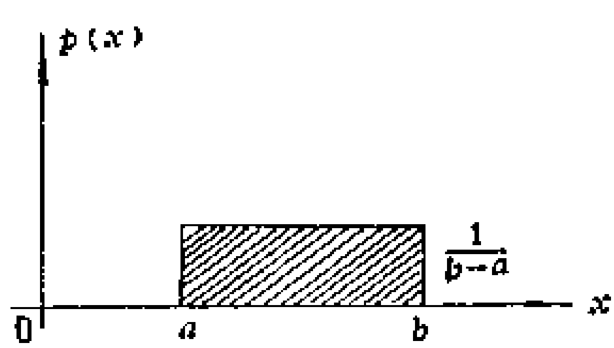


图 (3—1题)

概率$P(a\leqslant\xi\leqslant b)=$ "密度曲线与$[a,b]$区间之间所夹的面积"

$=\frac{1}{b-a}(b-a)=1$。

(2) (a)　$p\{a\leqslant\xi\leqslant c\}=\int_a^c p(x)dx=\int_a^c\frac{1}{b-a}dx$

$$= \frac{c-a}{b-a}$$

(b) $p(a \leqslant \xi \leqslant c) = \int_a^c p(x)dx = \int_a^b p(x)dx$

$$+ \int_b^c p(x)dx$$

$$= \int_a^b \frac{1}{b-a}dx + \int_b^c 0 \cdot dx = 1 + 0$$

$$= 1$$

2 (1) ∵ $\int_{-\infty}^{\infty} p(x)dx = 1$

由 $\int_{-\infty}^{\infty} p(x)dx = \int_0^1 axdx + \int_1^2 adx$

$$+ \int_2^3 (-ax + 3a)dx$$

$$= a \cdot \frac{x^2}{2}\Big|_0^1 + ax\Big|_1^2 + \left(-a \cdot \frac{x^2}{2} + 3ax\right)\Big|_2^3$$

$$= \frac{a}{2} + a + \left(-\frac{9}{2}a + 9a + 2a - 6a\right)$$

$$= \frac{3}{2}a + \frac{1}{2}a$$

$$= 2a$$

∴ $2a = 1$

∴ $a = \frac{1}{2}$.

(2) 设：A表示“三个观测值中正好一个大于1.5”的事件，A_i表示“观测值ξ_i大于1.5且其余两个观测值均小于等于1.5”的事件。($i = 1, 2, 3$)

$$P(A) = P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3)$$

$$= 3P(A_1) = 3P((\xi_1 > 1.5) \cdot (\xi_2 \leqslant 1.5) \cdot (\xi_3 \leqslant 1.5))$$

$$= 3P(\xi_1 > 1.5) \cdot [P(\xi_2 \leqslant 1.5)]^2$$

$$= 3 \cdot \int_{1.5}^{\infty} f(x)dx \cdot \left(\int_{-\infty}^{1.5} f(x)dx\right)^2$$

$$= 3\cdot\left[\int_{1.5}^{2}\frac{1}{2}dx + \int_{2}^{3}\frac{1}{2}(-x+3)dx\right]\cdot$$

$$\cdot\left(\int_{0}^{1}\frac{x}{2}dx + \int_{1}^{1.5}\frac{1}{2}dx\right)^2$$

$$= 3\cdot\left[\frac{1}{4}+\left(-\frac{9}{4}+\frac{4}{4}+\frac{3}{2}\right)\right]\cdot\left(\frac{1}{4}+\frac{1}{4}\right)^2$$

$$= 3\times\frac{1}{2}\times\frac{1}{4}$$

$$= \frac{3}{8}.$$

3　（1）∵　$P(\xi>1)=\frac{1}{3}$

而　$P(\xi>1)=\int_{1}^{a}\frac{1}{2a}dx = \frac{1}{2a}(a-1)$

于是有，$\frac{1}{2a}(a-1)=\frac{1}{3}$

∴　解得$a=3$。

（2）∵　$P(\xi>1)=\frac{1}{2}$

而　$P(\xi>1)=\int_{1}^{a}\frac{1}{2a}dx=\frac{1}{2a}(a-1)$

于是有　$\frac{1}{2a}(a-1)=\frac{1}{2}$，

即　　$a-1=a$

故，此题无解。

（3）∵　$P\left(\xi<\frac{1}{2}\right)=0.7$

而　$P\left(\xi<\frac{1}{2}\right)=\int_{-a}^{\frac{1}{2}}\frac{1}{2a}dx=\frac{\frac{1}{2}+a}{2a}$

于是有　$\frac{\frac{1}{2}+a}{2a}=\frac{7}{10}$

∴　解得　　$a=\frac{5}{4}$。

（4）$\because \quad P\left(\xi<\frac{1}{2}\right)=0.3$

而 $\quad P\left(\xi<\frac{1}{2}\right)=\int_{-a}^{\frac{1}{2}}\frac{1}{2a}dx=\frac{\frac{1}{2}+a}{2a}$

于是有，$\frac{\frac{1}{2}+a}{2a}=\frac{3}{10}$

$\therefore \quad a=-\frac{5}{4}$

由于 a 不能为负值，

故 a 不存在。

（5）$\because \quad P\{|\xi|<1\}=P\{|\xi|>1\}$

而 $\quad P(|\xi|<1)=P(-1<\xi<1)=\int_{-1}^{1}\frac{1}{2a}dx=\frac{1}{a}$

$$
\begin{aligned}
P\{|\xi|>1\}&=P((\xi<-1)\cup(\xi>1))\\
&=P(\xi<-1)+P(\xi>1)\\
&=\int_{-a}^{-1}\frac{1}{2a}dx+\int_{1}^{a}\frac{1}{2a}dx\\
&=\frac{1}{2a}(a-1+a-1)\\
&=\frac{a-1}{a}
\end{aligned}
$$

于是有 $\quad \frac{1}{a}=\frac{a-1}{a}(a\neq 0)$

$\therefore$ 解得 $a=2$

4 （1）设 $A=$ “$\xi\geqslant 200$小时”；

$B=$ “$\xi\geqslant 150$小时”。

$$
P(A|B)=\frac{P(AB)}{P(B)}=\frac{P((\xi\geqslant 200)\cdot(\xi\geqslant 150))}{P\{\xi\geqslant 150\}}
$$

$$
=\frac{P\{\xi\geqslant 200\}}{P\{\xi\geqslant 150\}}=\frac{\int_{200}^{\infty}\frac{100}{x^2}dx}{\int_{150}^{\infty}\frac{100}{x^2}dx}
$$

$$=\frac{-\frac{100}{x}\Big|_{200}^{\infty}}{-\frac{100}{x}\Big|_{150}^{\infty}}=\frac{\frac{1}{2}}{\frac{2}{3}}=\frac{3}{4}$$

（2） $P(〔\xi_1\leqslant150)\cdot(\xi_2>150)\cdot(\xi_3>150)〕\cup〔(\xi_1>150)\cdot$

$\cdot(\xi_2\leqslant150)\cdot(\xi_3>150)〕\cup〔(\xi_1>150)\cdot$

$\cdot(\xi_2>150)\cdot(\xi_3\leqslant150〕)$

$=3P(\xi_1\leqslant150\}\cdot〔P\{\xi_2>150)〕^2$

$=3\left(\int_{-\infty}^{150}f(x)dx\right)\cdot\left[\int_{150}^{\infty}f(x)dx\right]^2$

$=3\left(1-\int_{150}^{\infty}f(x)dx\right)\left[\int_{150}^{\infty}f(x)dx\right]^2$

$=3\left[1-\frac{2}{3}\right]\left[\frac{2}{3}\right]^2=3\times\frac{1}{3}\times\frac{4}{9}=\frac{4}{9}$.

（3）设有n个这种管子。

$P((\xi_1\geqslant150)\cdot(\xi_2\geqslant150)\cdots(\xi_n\geqslant150))\geqslant\frac{1}{2}$

$(P(\xi_1\geqslant150))^n\geqslant\frac{1}{2}$

$\left(\int_{150}^{\infty}\frac{100}{x^2}dx\right)^n\geqslant\frac{1}{2}$

$\left(\frac{2}{3}\right)^n\geqslant\frac{1}{2}$

显然$n=1$。

5　设ξ的分布密度$p(x)$为原点对称$p(x)=p(-x)$。

（1） $1=\int_{-\infty}^{\infty}p(x)dx=\int_{-\infty}^{0}p(x)dx+\int_{0}^{\infty}p(x)dx$

$=2\int_{-\infty}^{0}p(x)dx$

所以　　$\int_{-\infty}^{0}p(x)dx=F(0)=\frac{1}{2}$

（2） $p(\xi>a)=p(\xi\geqslant a)=1-p(\xi\leqslant a)=1-\int_{-\infty}^{a}p(x)dx$

$$=1-\left[\int_{-\infty}^{0} p(x)dx+\int_{0}^{a} p(x)dx\right]$$

$$=0.5-\int_{0}^{a} p(x)dx$$

(3) $F(-a)+F(a)=1$

$$F(-a)+F(a)=\int_{-\infty}^{-a} p(x)dx+\int_{-\infty}^{a} p(x)dx$$

$$=\int_{a}^{\infty} p(x)dx+\int_{-\infty}^{a} p(x)dx$$

$$=\int_{-\infty}^{\infty} p(x)dx=1$$

(4) $P(-a<\xi<a)=P(-a<\xi\leqslant a)=F(a)-F(-a)$

$$=F(a)-[1-F(a)]=2F(a)-1$$

(5) $P(|\xi|>a)=1-P(|\xi|\leqslant a)=1-P(-a\leqslant\xi\leqslant a)$

$$=1-[2F(a)-1]=2[1-F(a)]=2F(-a)$$

6 $$P\left\{\xi>b\,\middle|\,\xi<\frac{b}{2}\right\}=\frac{P\left\{(\xi>b)\cdot\left(\xi<\frac{b}{2}\right)\right\}}{P\left\{\xi<\frac{b}{2}\right\}}=\frac{P\left\{b<\xi<\frac{b}{2}\right\}}{P\left\{\xi<\frac{b}{2}\right\}}$$

$$=\frac{\int_{b}^{\frac{b}{2}}3x^2dx}{\int_{-1}^{\frac{b}{2}}3x^2dx}=\frac{x^3\Big|_{b}^{\frac{b}{2}}}{x^3\Big|_{-1}^{\frac{b}{2}}}$$

$$=\frac{\frac{b^3}{8}-b^3}{\frac{b^3}{8}+1}=\frac{b^3-8b^3}{b^3+8}$$

$$=\frac{-7b^3}{b^3+8}$$

7 证明：

(1) ∵ f，g 是同一区间 $[a,b]$ 上的密度函数

∴ 有 $\int_{a}^{b} f(x)dx=1$，$\int_{a}^{b} g(x)dx=1$。

而 $\int_a^b (f+g)dx = \int_a^b fdx + \int_a^b gdx$

$= 1+1=2$

故$f+g$不是$[a,b]$上的密度函数。

（2）∵ 1° 由$0<\beta<1$，f，$g\geqslant 0$

有$\beta f(x)+(1-\beta)g(x)\geqslant 0$且可积；

2° $\int_a^b [\beta f(x)+(1-\beta)g(x)]dx$

$=\beta\int_a^b f(x)dx+(1-\beta)\int_a^b g(x)dx$

$=\beta\times 1+(1-\beta)\times 1$

$=1$

∴ $\beta f+(1-\beta)g$是$[a,b]$上的密度函数。

8 题中给出的折线为一随机变量的分布密度，所以折线与 x 轴所夹面积S应等于 1，另一方面此三角形面积$S=\frac{1}{2}[b-(-a)]\cdot b$，

所以

$$\frac{1}{2}[b+a]\cdot b=1$$

解得 $a=\frac{2}{b}-b$

9 由于$y=ax+b(a\neq 0)$是线性函数，所以严格单调且可微。由定理3.3.1得

$$\varphi(y)=g(x)\left|\frac{dx}{dy}\right|=g\left(\frac{y-b}{a}\right)\left|\frac{1}{a}(y-b)'\right|$$

$$=\left|\frac{1}{a}\right|\cdot g\left(\frac{y-b}{a}\right).$$

10 $G(y)=P(\eta\leqslant y)=P(4-\xi^2\leqslant y)$

$=P(\xi^2\geqslant 4-y)$

$=P((\xi\leqslant-\sqrt{4-y})\cup(\xi\geqslant\sqrt{4-y}))$

$=P(\xi\leqslant-\sqrt{4-y})+P(\xi\geqslant\sqrt{4-y})$

$$=\int_{-1}^{-\sqrt{4-y}}\frac{1}{2}dx+\int_{\sqrt{4-y}}^{1}\frac{1}{2}dx$$

$$=\frac{1}{2}[(-\sqrt{4-y})+1+1-\sqrt{4-y}]$$

$$=1-\sqrt{4-y}$$

$\therefore\quad g(y)=G'(y)=\begin{cases}\dfrac{1}{2\sqrt{4-y}} & 3<y<4\\ 0 & 其它\end{cases}$

$\because\quad -1<x<1$

即　$0<|x|<1$

$\therefore\quad 0<\sqrt{4-y}<1$

$\quad 0<4-y<1$

解得$3<y<4$

作图略。

11　（1）由$y=x^3$，有$x=\sqrt[3]{y}$

由$x>0$知$y=x^3$是单调增函数

$\therefore$　由定理3.3.1

$$g(y)=f(x)\left|\frac{dx}{dy}\right|=e^{-y^{\frac{1}{3}}}\left(\frac{1}{3}y^{-\frac{2}{3}}\right)$$

$$=\frac{1}{3}y^{-\frac{2}{3}}e^{-y^{\frac{1}{3}}}$$

$\because\quad H(x)=x^3$是增函数

$\therefore\quad H(a)<y<H(b)$

即　$0<y<\infty$

（2）$z=\dfrac{3}{(x+1)^2}$，$x=\sqrt{\dfrac{3}{z}}-1$　　　（$\because x>0$）

而$H(x)$是单调减函数

由定理3.3.1

$$g(z)=f(x)\left|\frac{dx}{dy}\right|=e^{-\sqrt{\frac{3}{z}}+1}\cdot\left[\frac{1}{2\sqrt{\frac{3}{z}}}\left(-\frac{3}{z^2}\right)\right]$$

$$=\frac{\sqrt{3}}{2}\cdot z^{-\frac{3}{2}}e^{-\sqrt{\frac{3}{z}}+1}$$

$\because$ $H(x)$递减

$\therefore$ $H(b)<z<H(a)$

即 $0<z<3$

12 （1）设$y=x^{-1}$.

1° $y>0$时，$G(y)=P(\eta\leqslant y)=P\left(\frac{1}{\xi}\leqslant y\right)$

$$=P\left\{\left(\xi\geqslant\frac{1}{y}\right)\cup(\xi\leqslant 0)\right\}$$

$$=P\left(\xi\geqslant\frac{1}{y}\right)+P(\xi\leqslant 0)$$

$$=1-P\left(\xi<\frac{1}{y}\right)+P(\xi\leqslant 0)$$

$$=1-F(y^{-1})+F(0)$$

2° $y=0$时，$G(y)=P\left(\frac{1}{\xi}\leqslant y\right)=P(\xi\leqslant 0)=F(0)$

3° $y<0$时 $G(y)=P\left(\frac{1}{\xi}\leqslant y\right)=P(y^{-1}\leqslant\xi\leqslant 0)$

$$=F(0)-F(y^{-1})$$

$$\therefore\quad G(y)=\begin{cases}1-F(y^{-1})+F(0), & y>0;\\ F(0) & y=0;\\ F(0)-F(y^{-1}) & y<0.\end{cases}$$

$$g(y)=G'(y)=\begin{cases}y^{-2}\cdot p(y^{-1}) & y\neq 0\\ 0 & y=0\end{cases}$$

（2）设$y=|x|$

1° $y\geqslant 0$时，

$$G(y)=P(\zeta\leqslant y)=P(|\xi|\leqslant y)$$

$$=P(-y\leqslant\xi\leqslant y)$$

$$=F(y)-F(-y)$$

2° $y<0$时，

$$G(y)=0$$

$$\therefore\quad G(y)=\begin{cases}F(y)-F(-y) & y\geqslant 0;\\ 0 & y<0.\end{cases}$$

$$g(y)=\begin{cases}p(y)+p(-y) & y\geqslant 0;\\ 0 & y<0.\end{cases}$$

13　（ⅰ）$\eta=F(\xi)$是随机变量，事实上，η的的取值在区间$[0,1]$里，对任意$x\in[0,1]$，$(\eta<x)=(\xi<F^{-1}(x))\in$事件$\sigma$－域$\mathscr{F}$。而对任意$x>1$，$(\eta<x)$为必然事件，对任意$x<0$，$(\eta<x)$为不能事件。

（ⅱ）对任意$a,b\in[0,1]$

$$P(a<\eta<b)=P(F^{-1}(a)<\xi<F^{-1}(b))$$

$$=\int_{F^{-1}(a)}^{F^{-1}(b)}f(x)dx$$

$$=F(F^{-1}(b))-F(F^{-1}(a))$$

$$=b-a=\int_b^a dx$$

故η为在$[0,1]$上均匀分布的随机变量，其分布密度为

$$f_\eta(x)=\begin{cases}1 & x\in[0,1]\\ 0 & \text{其他}.\end{cases}$$

14　（1）设　$y=\cos x$。

当　$|y|\leqslant 1$时，

$$G(y)=P(\eta\leqslant y)=P(\cos\xi\leqslant y)$$

$$=\sum_{n=-\infty}^{\infty}P(\cos\xi\leqslant y,2n\pi<\xi\leqslant 2(n+1)\pi)$$

$$=\sum_{n=-\infty}^{\infty}P(2n\pi+\cos^{-1}y\leqslant\xi\leqslant 2(n+1)\pi-\cos^{-1}y)$$

$$=\sum_{n=-\infty}^{\infty}[F(2(n+1)\pi-\cos^{-1}y)-F(2n\pi+\cos^{-1}y)]$$

当　$|y|>1$时，显然$G(y)=0$。所以，

$G(y)$

$$=\begin{cases}\sum\limits_{n=-\infty}^{\infty}[F(2(n+1)\pi-\cos^{-1}y)-F(2n\pi+\cos^{-1}y)], & |y|\leqslant 1;\\ 0 & \text{其它.}\end{cases}$$

$g(y)$

$$=\begin{cases}\frac{1}{\sqrt{1-y^2}}\cdot\sum\limits_{n=-\infty}^{\infty}[p(2(n+1)\pi-\cos^{-1}y)+p(2n\pi+\cos^{-1}y)], & |y|\leqslant 1;\\ 0 & \text{其它.}\end{cases}$$

（2）设 $y=\operatorname{tg}x$.

$$G(y)=P(\zeta\leqslant y)=P(\operatorname{tg}\xi\leqslant y)$$

$$=\sum_{n=-\infty}^{\infty}P\left(\operatorname{tg}\xi\leqslant y,\ -\frac{\pi}{2}+n\pi<\xi\leqslant\frac{\pi}{2}+n\pi\right)$$

$$=\sum_{n=-\infty}^{\infty}P\left(-\frac{\pi}{2}+n\pi<\xi\leqslant\operatorname{tg}^{-1}y+n\pi\right)$$

$$=\sum_{n=-\infty}^{\infty}\left[F(\operatorname{tg}^{-1}y+n\pi)-F\left(-\frac{\pi}{2}+n\pi\right)\right]$$

$$=\sum_{n=-\infty}^{\infty}\int_{-\frac{\pi}{2}+n\pi}^{\operatorname{tg}^{-1}y+n\pi}p(u)du$$

$$g(y)=G'(y)=\frac{1}{1+y^2}\cdot\sum_{n=-\infty}^{\infty}p(\operatorname{tg}^{-1}y+n\pi)$$

15 $\quad P(\eta\leqslant k)=P(1-\xi\leqslant k)=P(\xi\geqslant 1-k)=0.25$

∵ $\quad P(\eta\leqslant k)+P(\xi\leqslant 0.29)=0.25+0.75=1$

∴ $\quad P(\xi>1-k)+P(\xi\leqslant 0.29)=1$

因此，$1-k=0.29$

$\qquad k=0.71$

16 首先确定 c 值.

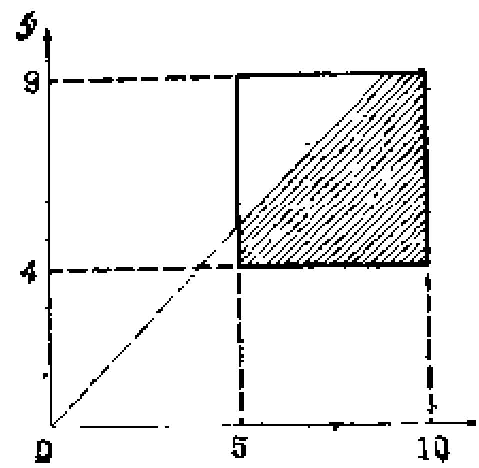
$$1=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}p(x,y)\,dxdy=\int_{4}^{9}\int_{5}^{10}c\,dxdy=c5^2$$

所以 $\qquad c=5^{-2}$

$$P(\xi_1 \geqslant \xi_2) = \iint\limits_{\Delta} p(x,y)\,dxdy$$

其中区域$\Delta=\{(x,y):\ x\geqslant y\}$，图(3—16)中的阴影部分为区域 Δ 里使 $p(x,y)$取值为 c 的部分，所以概率

$$P(\xi_1 \geqslant \xi_2) = 5^{-2}\left[5^2 - \frac{1}{2}4^2\right] = \frac{17}{25}$$



图(3—16)题

17　(1) $\because$ $\displaystyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dx\cdot dy = \int_0^2\int_{-x}^{x} kx(x-y)dydx$

$$= \int_0^2 kx\left[xy - \frac{y^2}{2}\right]_{-x}^{x} dx$$

$$= \int_0^2 kx\left[\left(x^2 - \frac{x^2}{2}\right) - \left(-x^2 - \frac{x^2}{2}\right)\right]dx$$

$$= \int_0^2 kx(x^2 + x^2)dx$$

$$= 2k\int_0^2 x^3 dx$$

$$= 2k\frac{x^4}{4}\Big|_0^2$$

$$= 8k$$

$$= 1$$

$\therefore\quad k = \dfrac{1}{8}$

(2) $\displaystyle h_1(x) = \int_{-x}^{x} f(x,y)dy = \int_{-x}^{x}\frac{1}{8}(x^2 - xy)dy$

$$= \frac{1}{8}\left(x^2y - \frac{xy^2}{2}\right)\Big|_{-x}^{x}$$

$$= \frac{1}{8}(2x^3 - 0)$$

$$= \frac{x^3}{4} \qquad 0 < x < 2;$$

$y \geq 0$时，

$$h_2(y) = \int_y^2 f(x, y)dx = \int_y^2 \frac{1}{8}x(x-y)dx$$

$$= \frac{1}{8}\left(\frac{x^3}{3} - \frac{x^2}{2}y\right)\Big|_y^2 = \frac{1}{8}\left(\frac{8}{3} - \frac{y^3}{3} - 2y + \frac{y^3}{2}\right)$$

$$= \frac{1}{8}\left(\frac{8}{3} - 2y + \frac{y^3}{6}\right)$$

$$= \frac{1}{3} - \frac{y}{4} + \frac{y^3}{48}$$

$y < 0$时，

$$h_2(y) = \int_{-y}^2 f(x, y)dx = \int_{-y}^2 \frac{1}{8}x(x-y)dx$$

$$= \frac{1}{8}\left(\frac{x^3}{3} - \frac{x^2 y}{2}\right)\Big|_{-y}^2 = \frac{1}{8}\left(\frac{8}{3} + \frac{y^3}{3} - 2y + \frac{y^3}{2}\right)$$

$$= \frac{1}{8}\left(\frac{8}{3} - 2y + \frac{5}{6}y^3\right)$$

$$= \frac{1}{3} - \frac{y}{4} + \frac{5y^3}{48}$$

$$\therefore \quad h_2(y) = \begin{cases} \frac{1}{3} - \frac{y}{4} + \frac{y^3}{48} & 0 \leq y < 2 \\ \frac{1}{3} - \frac{y}{4} + \frac{5y^3}{48} & -2 < y < 0 \\ 0 & \text{其它.} \end{cases}$$

18　ξ的边缘分布密度为

$$h_1(x) = \int_0^2 f(x, y)dy = \int_0^2 \left(x^2 + \frac{xy}{3}\right)dy$$

$$= \left(x^2 y + \frac{xy^2}{6}\right)\Big|_0^2 = 2x^2 + \frac{2x}{3}$$

$$= 2\left(x^2 + \frac{1}{3}x\right) \qquad 0 < x < 1;$$

η的边缘分布密度为

$$h_2(y)=\int_0^1 f(x,y)\,dx=\int_0^1\left(x^2+\frac{xy}{3}\right)dx$$

$$=\left(\frac{x^3}{3}+\frac{x^2y}{6}\right)\Big|_0^1=\frac{1}{3}+\frac{y}{6}$$

$$=\frac{1}{3}\left(1+\frac{1}{2}y\right)\qquad 0<y<2.$$

（1） $$P\left(\xi>\frac{1}{2}\right)=\int_{\frac{1}{2}}^1 h_1(x)dx=\int_{\frac{1}{2}}^1 2\left(x^2+\frac{x}{3}\right)dx$$

$$=2\left(\frac{x^3}{3}+\frac{x^2}{6}\right)\Big|_{\frac{1}{2}}^1$$

$$=2\left[\left(\frac{1}{3}+\frac{1}{6}\right)-\frac{1}{24}-\frac{1}{24}\right]$$

$$=2\left(\frac{3}{6}-\frac{1}{12}\right)$$

$$=2\times\frac{5}{12}$$

$$=\frac{5}{6}$$

（2） $$P(\eta<\xi)=\int_0^1\int_0^x f(x,y)dydx$$

$$=\int_0^1\int_0^x\left(x^2+\frac{xy}{3}\right)dydx$$

$$=\int_0^1\left(x^2y+\frac{xy^2}{6}\right)\Big|_0^x dx$$

$$=\int_0^1\left(x^3+\frac{x^3}{6}\right)dx$$

$$=\int_0^1\frac{7}{6}x^3dx$$

$$=\frac{7}{6}\int_0^1 x^3dx$$

$$= \frac{7}{6} \cdot \frac{x^4}{4}\Big|_0^1$$

$$= \frac{7}{24}$$

（3） $$P\left(\eta<\frac{1}{2}\Big|\xi<\frac{1}{2}\right)=\frac{P\left\{\left(\eta<\frac{1}{2}\right)\cdot\left(\xi<\frac{1}{2}\right)\right\}}{P\left(\xi<\frac{1}{2}\right)}$$

$$=\frac{\int_0^{\frac{1}{2}}\int_0^{\frac{1}{2}}\left(x^2+\frac{xy}{3}\right)dxdy}{\int_0^{\frac{1}{2}}2\left(x^2+\frac{x}{3}\right)dx}$$

$$=\frac{\int_0^{\frac{1}{2}}2\left(\frac{x^3}{3}+\frac{x^2y}{6}\right)\Big|_0^{\frac{1}{2}}dy}{\left(\frac{2x^2}{3}+\frac{2x^2}{6}\right)\Big|_0^{\frac{1}{2}}}$$

$$=\frac{\int_0^{\frac{1}{2}}\left(\frac{1}{24}+\frac{y}{24}\right)dy}{\frac{1}{12}+\frac{1}{12}}$$

$$=\frac{6}{24}\int_0^{\frac{1}{2}}(1+y)dy$$

$$=\frac{1}{4}\left(y+\frac{y^2}{2}\right)\Big|_0^{\frac{1}{2}}$$

$$=\frac{1}{4}\left(\frac{1}{2}+\frac{1}{8}\right)$$

$$=\frac{5}{32}$$

19　设 ξ_k^*的分布函数为$F_k(x)$，则

$$F_k(x)=P(\xi_k^*\leqslant x)$$

$$=P(\xi_1,\xi_2,\cdots,\xi_n\text{中至少有 }k\text{ 个小于等于 }x)$$

设 A_i = "$\xi_1, \xi_2, \cdots, \xi_n$中恰有 i 个小于等于 x"，则

$$F_k(x) = P\left(\bigcup_{i=k}^{n} A_i\right)$$

$$= \sum_{i=k}^{n} P(A_i)$$

$$= \sum_{i=k}^{n} \binom{n}{i} (F(x))^i (1 - F(x))^{n-i}$$

20　ξ的分布密度 $p_1(x) = \begin{cases} 1 & 0 \leqslant x \leqslant 1 \\ 0 & \text{其它} \end{cases}$

η的分布密度 $p_2(x) = \begin{cases} \frac{1}{2} & 0 \leqslant x \leqslant 2 \\ 0 & \text{其它} \end{cases}$

则$Z = \xi + \eta$的分布密度

$$p_Z(z) = \int_{-\infty}^{\infty} p_1(x) p(z - x) dx$$

上式中被积函数在区域$0 \leqslant x \leqslant 1$，$0 \leqslant z - x \leqslant 2$里非零，此区域形状见图（3—20题）.

当　$z > 3$　或　$z < 0$时，

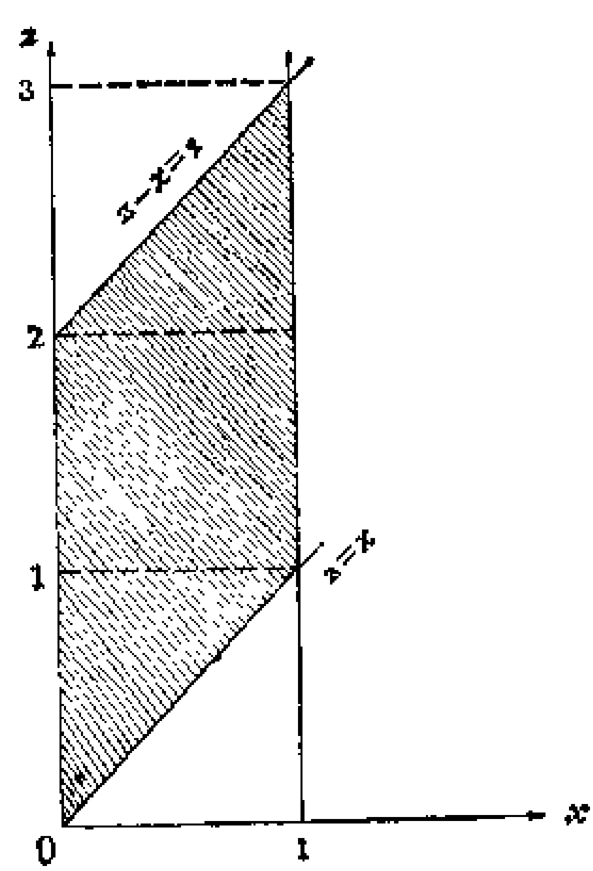
$p_Z(z) = 0$.

当　$2 \leqslant z \leqslant 3$时，

$$p_Z(z) = \int_{z-2}^{1} \frac{1}{2} dx$$

$$= \frac{1}{2} x \Big|_{z-2}^{1}$$

$$= \frac{1}{2}(3 - z)$$



图（3—20题）

当　$1 \leqslant z \leqslant 2$时，

$$p_Z(z)=\int_0^1\frac{1}{2}dx=\frac{1}{2}$$

当 $0\leqslant z\leqslant 1$时，

$$p_Z(z)=\int_0^z\frac{1}{2}dx=\frac{z}{2}$$

21 设 $\eta=\dfrac{\xi_1}{\xi_2}$的分布密度为$g(x)$，则

$$g(x)=\int_{-\infty}^{\infty}f(xz)f(z)|z|dz.$$

$$=\begin{cases}\displaystyle\int_{1000}^{\infty}z\frac{1000}{(xz)^2}\cdot\frac{1000}{z^2}dz & x\geqslant 1\\ \displaystyle\int_{\frac{1000}{x}}^{\infty}z\frac{1000}{(xz)^2}\cdot\frac{1000}{z^2}dz & 0<x<1\\ 0 & \text{其它}\end{cases}$$

$$=\begin{cases}\dfrac{1}{2x^2} & x\geqslant 1\\ \dfrac{1}{2} & 0<x<1\\ 0 & \text{其它}\end{cases}$$

22 ξ_1的边缘分布密度为

$$h_1(x_1)=\int_0^{\infty}\frac{1}{\Gamma(k_1)\cdot\Gamma(k_2)}x_1^{k_1-1}(x_2-x_1)^{k_2-1}e^{-x_2}dx_2$$

$$=\frac{x_1^{k_1-1}}{\Gamma(k_1)\cdot\Gamma(k_2)}\cdot\int_{x_1}^{\infty}(x_2-x_1)^{k_2-1}\cdot e^{-x_2}dx_2$$

$$=\frac{x_1^{k_1-1}}{\Gamma(k_1)\cdot\Gamma(k_2)}\int_0^{\infty}t^{k_2-1}\cdot e^{-(t+x_1)}dt\quad (t=x_2-x_1)$$

$$=\frac{x_1^{k_1-1}\cdot e^{-x_1}}{\Gamma(k_1)\cdot\Gamma(k_2)}\int_0^{\infty}t^{k_2-1}\cdot e^{-t}dt$$

$$=\frac{x_1^{k_1-1}\cdot e^{-x_1}}{\Gamma(k_1)\cdot\Gamma(k_2)}\cdot\Gamma(k_2)$$

$$=\frac{1}{\Gamma(k_1)}x_1^{k_1-1}e^{-x_1}\qquad (0<x_1<\infty)$$

ξ_2的边缘分布密度为

$$h_2(x_2)=\int_0^{x_2}\frac{1}{\Gamma(k_1)\cdot\Gamma(k_2)}x_1^{k_1-1}(x_2-x_1)^{k_2-1}e^{-x_2}dx_1$$

$$=\frac{e^{-x_2}}{\Gamma(k_1)\cdot\Gamma(k_2)}\int_0^{x_2}x_1^{k_1-1}(x_2-x_1)^{k_2-1}dx_1$$

$$=\frac{e^{-x_2}}{\Gamma(k_1)\cdot\Gamma(k_2)}\int_0^{x_2}x_2^{k_1}\left(\frac{x_1}{x_2}\right)^{k_1-1}$$

$$\left[x_2^{k_2-1}\left(1-\frac{x_1}{x_2}\right)^{k_2-1}\right]d\left(\frac{x_1}{x_2}\right)$$

$$=\frac{e^{-x_2}\cdot x_2^{k_1+k_2-1}}{\Gamma(k_1)\cdot\Gamma(k_2)}\int_0^1 t^{k_1-1}(1-t)^{k_2-1}dt$$

$$(\because x_1\leqslant x_2)$$

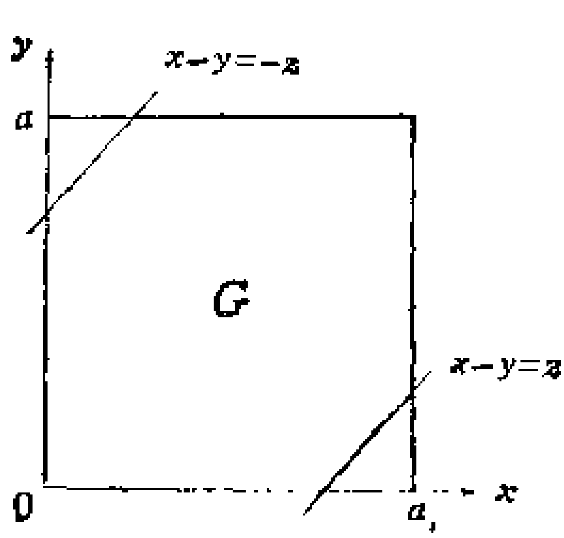
$$=\frac{e^{-x_2}\cdot x_2^{k_1+k_2-1}}{\Gamma(k_1)\cdot\Gamma(k_2)}\cdot\frac{\Gamma(k_1)\cdot\Gamma(k_2)}{\Gamma(k_1+k_2)}$$

$$=\frac{1}{\Gamma(k_1+k_2)}x_2^{k_1+k_2-1}\cdot e^{-x_2}\qquad(0<x_2<\infty)$$

23　设ξ,η是所投的两点的坐标，则显然其联合分布密度为

$$f_{\xi\eta}(x,y)=\begin{cases}0 & (x,y)\overline{\in}[0,a]\times[0,a]\\ \dfrac{1}{a^2} & (x,y)\in[0,a]\times[0,a].\end{cases}$$

两点的距离为$\zeta\triangleq|\xi-\eta|$，则其分布函数为



图(3—21)

$$F_\zeta(z)=P(|\xi-\eta|\leqslant z)$$

$$=\begin{cases}0 & z\leqslant 0\\ \iint\limits_{|x-y|\leqslant z}\frac{1}{a^2}dxdy & 0\leqslant z<a\\ 1 & z\geqslant a\end{cases}$$

面积分 $\iint\limits_{|x-y|\leqslant z}\frac{1}{a^2}dxdy$

$$=\frac{1}{a^2}\cdot(G\text{的面积})$$

$$=\frac{1}{a^2}(a^2-(a-z)^2)$$

$$=\frac{(2az-z^2)}{a^2}$$

24　先计算I^2的分布，其分布函数

$$F_{I^2}(x)=P(I^2\leqslant x)=P(-\sqrt{x}\leqslant I\leqslant\sqrt{x})$$

$$=\int_0^{\sqrt{x}}6i(1-i)di=3x-2x^{\frac{3}{2}},\quad 0\leqslant x\leqslant 1$$

故I^2的分布密度

$$h(x)=F'_{I^2}(x)=\begin{cases}3(1-\sqrt{x}) & 0\leqslant x\leqslant 1\\ 0 & \text{其它}\end{cases}$$

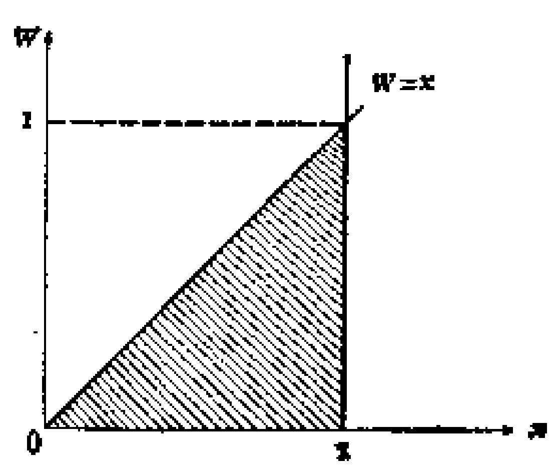
下面计算$W=I^2R$的分布密度．由定理3.3.4，W的密度函数

$$p_W(\omega)=\int_{-\infty}^{\infty}h(x)g\left(\frac{\omega}{x}\right)\frac{1}{|x|}dx$$

被积函数在区域$0\leqslant x\leqslant 1$，

$0\leqslant\frac{\omega}{x}\leqslant 1$内非零，此区域如

图(3—24题)的阴影部分．



图(3—24题)

当　$\omega>1$或$\omega<0$时

$$p_W(\omega)=0.$$

当　$0\leqslant\omega\leqslant 1$时，

$$p_{\overline{W}}(\omega)=\int_{\omega}^{1}3(1-\sqrt{x})2\left(\frac{\omega}{x}\right)\frac{1}{|x|}dx$$

$$=6\omega\int_{\omega}^{1}(x^{-2}-x^{-\frac{3}{2}})dx$$

$$=6\omega(1+\omega^{-1}-2\omega^{-\frac{1}{2}})$$

$$=6\omega-12\omega^{\frac{1}{2}}+6$$

25 设 ξ 为乘客等车的时间.

∵ 乘客在[0, 3]区间里都是等可能到达的,

∴ ξ 服从均匀分布. 即其分布密度为

$$p(x)=\begin{cases}\frac{1}{3} & 0<x<3\\ 0 & \text{其它}\end{cases}$$

于是有

$$E\xi=\int_{0}^{3}x\cdot\frac{1}{3}dx=\frac{1}{3}\cdot\frac{x^2}{2}\Big|_0^3$$

$$=\frac{9}{6}=1.5$$

26 (1) $\int_{2000}^{1\,000}k/x^2dx=-\frac{k}{x}\Big|_{2000}^{10000}=-\frac{k}{10000}+\frac{k}{2000}=1$

$$k\left(\frac{5-1}{10000}\right)=1$$

$$k=2500$$

(2) $\int_{2000}^{10000}k/x^3dx=k\int_{2000}^{10000}x^{-3}dx=k\left(\frac{x^{-2}}{-2}\right)\Big|_{2000}^{10000}$

$$=\frac{k}{2}\left(\frac{-1}{(10000)^2}+\frac{1}{(2000)^2}\right)$$

$$=\frac{k}{2}\left(\frac{25-1}{10^8}\right)$$

$$=k\cdot\frac{12}{10^8}=1$$

$$k=\frac{10^8}{12}\doteq 8,333,333$$

(3) $\int_{2000}^{10000} \frac{k}{x^n} dx = k\int_{2000}^{10000} x^{-n} dx = k\frac{x^{1-n}}{1-n}\Bigg|_{2000}^{10000}$

$$= k\left(\frac{(10000)^{1-n} - (2000)^{1-n}}{1-n}\right) = 1$$

$$k = \frac{1-n}{2000^{1-n}(5^{1-n}-1)}$$

(4) $E\xi_1 = \int_{2000}^{10000} x\cdot\frac{k}{x^2}dx = k\int_{2000}^{10000}\frac{1}{x}dx$

$$= k(\ln 10000 - \ln 2000) = k\cdot\ln\frac{10000}{2000}$$

$$= k\cdot\ln 5 \doteq 2500\times 1.6 = 4000.$$

$$E\xi_2 = k\int_{2000}^{10000} x/x^3 dx = k\int_{2000}^{10000}\frac{1}{x^2}dx = k\cdot\left(-\frac{1}{x}\right)\Bigg|_{2000}^{10000}$$

$$= k\left(\frac{4}{10000}\right) = k\cdot\frac{1}{2500} = \frac{8,333,333}{2500}$$

$$\doteq 3333.3$$

$$E\xi_3 = k\int_{2000}^{10000} x/x^n dx = k\int_{2000}^{10000} x^{1-n}dx$$

$$= k\cdot\frac{x^{2-n}}{2-n}\Bigg|_{2000}^{10000}$$

$$= k\cdot\frac{(10000)^{2-n} - (2000)^{2-n}}{2-n}$$

$$= k\cdot\frac{2000^{2-n}(5^{2-n}-1)}{2-n}$$

$$= \frac{1-n}{2000^{1-n}(5^{1-n}-1)}\cdot\frac{2000^{2-n}(5^{2-n}-1)}{2-n}$$

$$= 2000\cdot\frac{n-1}{n-2}\cdot\frac{5^{2-n}-1}{5^{1-n}-1}$$

27 $P(\xi\leqslant 0.9) = \int_0^{0.9} e^{-x}dx = -e^{-x}\Big|_0^{0.9} = -e^{-0.9}+1$

$$= 1 - \frac{1}{e^{0.9}} \doteq 1 - 0.407 = 0.593$$

$$P(\xi>0.9)=0.407$$

∴ 平均利润为 $[0.593\times(-2.00)]+(0.407\times3.00)$

$$=1.221-1.186$$

$$=0.035\text{（元）}$$

28 （1）∵ $\int_0^{\infty} c_1 x^{\alpha} e^{-\beta x}dx=1$

∴ $c_1=\dfrac{1}{\int_0^{\infty} x^{\alpha} e^{-\beta x}dx}$

设 $\beta x=t$，$x=\dfrac{t}{\beta}$

$$c_1=\frac{1}{\int_0^{\infty}\left(\frac{t}{\beta}\right)^{\alpha} e^{-t}\left(\frac{1}{\beta}\right)dt}$$

$$=\frac{\beta^{\alpha+1}}{\int_0^{\infty} t^{\alpha} e^{-t}dt}$$

$$=\frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)}$$

同理，$c_2=\dfrac{\beta^{\gamma+1}}{\Gamma(\gamma+1)}$

（3） $E\xi=\int_0^{\infty} x\cdot\dfrac{\beta^{\alpha+1}}{\Gamma(\alpha+1)}\cdot x^{\alpha}\cdot e^{-\beta x}dx$

$$=\frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)}\int_0^{\infty} x^{\alpha+1} e^{-\beta x}dx$$

$$=\frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)}\cdot\frac{\Gamma(\alpha+2)}{\beta^{\alpha+2}}=\frac{\alpha+1}{\beta}$$

$$E\xi^2=\int_0^{\infty} x^2\frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)}x^{\alpha}e^{-\beta x}dx$$

$$=\frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)}\int_0^{\infty} x^{\alpha+2}\cdot e^{-\beta x}dx$$

$$= \frac{\beta^{\alpha+1}}{\Gamma(\alpha+1)} \cdot \frac{\Gamma(\alpha+3)}{\beta^{\alpha+3}}$$

$$= \frac{(\alpha+2)(\alpha+1)}{\beta^2}$$

$$\therefore \quad D\xi = E\xi^2 - (E\xi)^2 = \frac{(\alpha+2)(\alpha+1)}{\beta^2} - \left(\frac{\alpha+1}{\beta}\right)^2$$

$$= \frac{(\alpha+1)[(\alpha+2)-(\alpha+1)]}{\beta^2}$$

$$= \frac{\alpha+1}{\beta^2}$$

同理，$E\eta = \int_0^{\infty} \frac{\beta^{\gamma+1}}{\Gamma(\gamma+1)} x^{\gamma+1} e^{-\beta x} dx$

$$= \frac{\beta^{\gamma+1}}{\Gamma(\gamma+1)} \cdot \frac{\Gamma(\gamma+2)}{\beta^{\gamma+2}} = \frac{\gamma+1}{\beta}$$

$$E\eta^2 = \int_0^{\infty} \frac{\beta^{\gamma+1}}{\Gamma(\gamma+1)} \cdot x^{\gamma+2} e^{-\beta x} dx$$

$$= \frac{\beta^{\gamma+1}}{\Gamma(\gamma+1)} \cdot \frac{\Gamma(\gamma+3)}{\beta^{\gamma+3}} = \frac{(\gamma+2)(\gamma+1)}{\beta^2}$$

$$\therefore \quad D\eta = E\eta^2 - (E\eta)^2 = \frac{\gamma+1}{\beta^2}$$

故 $E(\xi+\eta) = E\xi + E\eta = \frac{\alpha+\gamma+2}{\beta}$

$$D(\xi+\eta) = D\xi + D\eta = \frac{\alpha+\gamma+2}{\beta^2}$$

（2）设$\xi+\eta$的分布密度为$g(x)$，则

$$g(x) = \int_{-\infty}^{\infty} p_\xi(x-z) p_\eta(z) dz$$

$$= \int_0^x c_1 (x-z)^\alpha e^{-\beta(x-z)} \cdot c_2 z^\gamma e^{-\beta z} dz, \qquad (0 < x)$$

$$= c_1 \cdot c_2 e^{-\beta x} \int_0^x (x-z)^\alpha \cdot z^\gamma dz$$

$$= c_1 \cdot c_2 \cdot e^{-\beta x} \cdot x^{\alpha+\gamma+1} \int_0^1 (1-u)^\alpha u^\gamma dx \left(令 u = \frac{z}{x}\right)$$

$$= c_1 \cdot c_2 \cdot e^{-\beta x} \cdot x^{\alpha+\nu+1} \cdot \frac{\Gamma(\alpha+1)\Gamma(\nu+1)}{\Gamma(\alpha+\nu+2)}$$

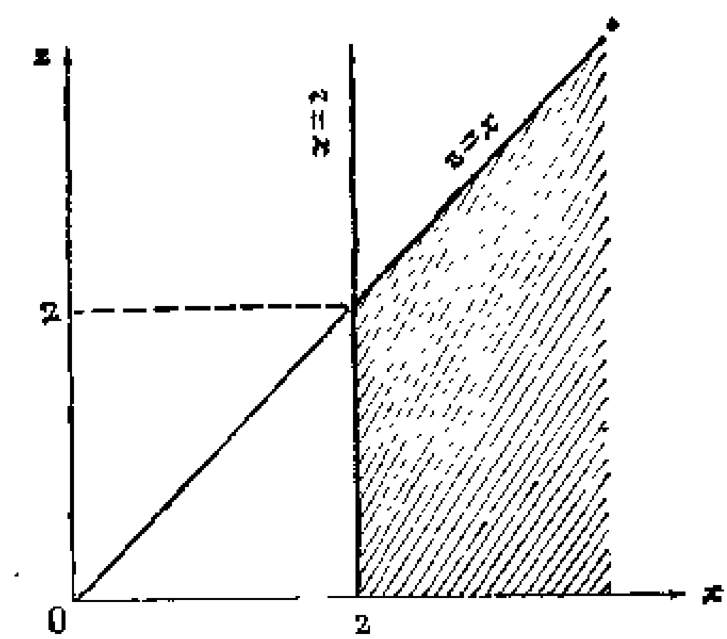
$$= \frac{\beta^{\alpha+\nu+2}}{\Gamma(\alpha+\nu+2)} x^{\alpha+\nu+1} e^{-\beta x}$$

29　（1）由定理3.3.4，ζ的密度函数

$$p_\zeta(z) = \int_{-\infty}^{\infty} f(x) g\left(\frac{z}{x}\right) \frac{1}{|x|} dx$$

被积函数的非零区域为 $x>2$，$0<\frac{z}{x}<1$，如图（3—29题）中的阴影部分。显然，当$z<0$时，$p_\zeta(z)=0$。

当　$0\leqslant z\leqslant 2$时，



图（3—29题）

$$p_\zeta(z) = \int_2^{\infty} \frac{8}{x^3} \cdot 2\frac{z}{x} \frac{1}{x} dx$$

$$= 16z\left(-\frac{1}{4}x^{-4}\right)\Big|_2^{\infty}$$

$$= \frac{z}{4}$$

当　$z>2$时

$$p_\zeta(z) = \int_z^{\infty} \frac{8}{x^3} \cdot 2\frac{z}{x} \frac{1}{x} dx = 4z^{-3}$$

综上所述，$p_\zeta(z)$的表达式为

$$p_\zeta(z)=\begin{cases}4z^{-3} & z>2\\ \dfrac{z}{4} & 0\leqslant z\leqslant 2\\ 0 & \text{其它}\end{cases}$$

（2）（i）$E\zeta=\int_{-\infty}^{\infty}z\cdot p_\zeta(z)dz$

$$=\int_2^{\infty}z\cdot 4z^{-3}dz+\int_0^2 z\cdot\frac{z}{4}dz$$

$$=2+\frac{2}{3}=2\frac{2}{3}$$

（ii） 设(ξ,η)的联合分布密度为 $p(x,y)$，则 $p(x,y)=f(x)\cdot g(y)$。这样，

$$E\zeta=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xyf(x)\cdot g(y)dxdy$$

$$=\int_2^{\infty}x\cdot\frac{8}{x^3}\ dx\cdot\int_0^1 y\cdot 2ydy$$

$$=8(-x^{-1})\Big|_2^{\infty}\cdot(2/3y^3)\Big|_0^1=4\cdot\frac{2}{3}=2\frac{2}{3}$$

30 $$E\xi=\int_0^{\infty}x\cdot\frac{1}{x\beta\sqrt{2\pi}}\cdot e^{-\frac{1}{2\beta^2}(\ln x-\alpha)^2}dx$$

$$=\frac{1}{\beta\sqrt{2\pi}}\int_0^{\infty}e^{-\frac{1}{2\beta^2}(\ln x-\alpha)^2}dx$$

设 $y=\ln x-\alpha$

$x=e^{y+\alpha}$

$dx=e^{y+\alpha}dy$

且$x\to 0$时，$y\to-\infty$

∴ 原式$=\dfrac{1}{\beta\sqrt{2\pi}}\displaystyle\int_{-\infty}^{\infty}e^{-\frac{1}{2\beta^2}y^2}\cdot e^{y+\alpha}dy$

$$=\frac{e^{\frac{\beta^2}{2}+\alpha}}{\beta\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{1}{2\beta^2}(y^2-2\beta^2y+\beta^4)}dy$$

$$=\frac{e^{\alpha+\frac{\beta^2}{2}}}{\beta\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{1}{2\beta^2}(y-\beta^2)^2}dy$$

又设$z=y-\beta^2$

$dy=dz$

$$\therefore\quad 原式=\frac{e^{\alpha+\frac{\beta^2}{2}}}{\beta\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{z^2}{2\beta^2}}\ dz$$

$$由于\quad \frac{1}{\beta\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{z^2}{2\beta^2}}\ dz=1$$

所以，原式$=e^{\alpha+\frac{\beta^2}{2}}$

$$E\xi^2=\int_0^{\infty}x^2\frac{1}{x\beta\sqrt{2\pi}}e^{-\frac{1}{2\beta^2}(\ln x-\alpha)^2}dx$$

$$=\frac{1}{\beta\sqrt{2\pi}}\int_0^{\infty}xe^{-\frac{1}{2\beta^2}(\ln x-\alpha)^2}dx$$

$$=\frac{1}{\beta\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{y+\alpha}\cdot e^{-\frac{y^2}{2\beta^2}}\cdot e^{y+\alpha}dy$$

$$=\frac{e^{2\alpha}}{\beta\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{y^2}{2\beta^2}+2y}\ dy$$

$$=\frac{e^{2\alpha+2\beta^2}}{\beta\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{1}{2\beta^2}(y^2-4\beta^2y+4\beta^4)}dy$$

$$=\frac{e^{2\alpha+2\beta^2}}{\beta\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{1}{2\beta^2}(y-2\beta^2)^2}dy$$

设　$z=y-2\beta^2$

$dz=dy$

$$\therefore \quad 原式=\frac{e^{2\alpha+2\beta^2}}{\beta\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{z^2}{2\beta^2}}dz$$

$$由于\frac{1}{\beta\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{z^2}{2\beta^2}}dz=1$$

所以，原式$=e^{2(\alpha+\beta^2)}$

$$\therefore \quad D\xi=E\xi^2-(E\xi)^2=e^{2(\alpha+\beta^2)}-(e^{\alpha+\frac{\beta^2}{2}})^2$$

$$=e^{2\alpha+2\beta^2}-e^{2\alpha+\beta^2}=e^{2\alpha+\beta^2}(e^{\beta^2}-1)$$

31　设 $F(x)$, $p(x)$分别表示$[0,a]$上均匀分布随机变量的分布函数和分布密度，则ξ_1^*, ξ_n^*的分布密度可分别表为

$$g_1(x)=n[1-F(x)]^{n-1}p(x)=n\left(1-\frac{x}{a}\right)^{n-1}\frac{1}{a}$$

$$(0<x<a)$$

$$g_n(x)=n[F(x)]^{n-1}p(x)=n\left(\frac{x}{a}\right)^{n-1}\frac{1}{a}\qquad(0<x<a)$$

于是

$$E\xi_1^*=\int_0^a x\cdot n\left(1-\frac{x}{a}\right)^{n-1}\frac{1}{a}dx$$

$$\because \quad 1=\int_0^a n\left(1-\frac{x}{a}\right)^{n-1}\frac{1}{a}dx$$

$$\therefore \quad 1-\frac{E\xi_1^*}{a}=\int_0^a(1-x/a)n\left(1-\frac{x}{a}\right)^{n-1}\frac{1}{a}dx$$

$$=\int_0^a n\left(1-\frac{x}{a}\right)^n\frac{1}{a}dx$$

$$=-\frac{n}{a^{n+1}}(a-x)^{n+1}\frac{1}{n+1}\Big|_0^a=\frac{n}{n+1}$$

$$\therefore \quad E\xi_1^*=\left(1-\frac{n}{n+1}\right)a=\frac{a}{n+1}$$

$$1-\frac{E\xi_1^{*2}}{a^2}=\int_0^a\left(1+\frac{x}{a}\right)n\left(1-\frac{x}{a}\right)^n\frac{1}{a}dx$$

$$-1+\frac{E\xi^{*2}}{a^2}=-2\int_0^a n\left(1-\frac{x}{a}\right)^n\frac{1}{a}dx+$$

$$+\int_0^a n\left(1-\frac{x}{a}\right)^n\frac{1}{a}dx-\int_0^a\frac{x}{a}\,n\left(1-\frac{x}{a}\right)^n\frac{1}{a}dx$$

$$\frac{E\xi_1^{*2}}{a^2}-1=-2\frac{n}{n+1}+\int_0^a n\left(1-\frac{x}{a}\right)^{n+1}\frac{1}{a}dx$$

$$=-2\frac{n}{n+1}+\frac{n}{n+2}$$

$$\therefore\quad E\xi_1^{*2}=\left(1-2\frac{n}{n+1}+\frac{n}{n+2}\right)a^2=\frac{2a^2}{(n+1)(n+2)}$$

$$D\xi_1^*=\left(1-2\frac{n}{n+1}+\frac{n}{n+2}\right)a^2-\left(1-\frac{n}{n+1}\right)^2a^2$$

$$=\left[\frac{n}{n+2}-\left(\frac{n}{n+1}\right)^2\right]a^2=\frac{na^2}{(n+1)^2(n+2)}$$

下面求ξ_n^*的均值和方差。

$$E\xi_n^*=\int_0^a xn\left(\frac{x}{a}\right)^{n-1}\frac{1}{a}dx$$

$$=\frac{n}{a^n}\int_0^a x^n dx=\frac{n}{n+1}a$$

$$E\xi_n^{*2}=\int_0^a x^2n\left(\frac{x}{a}\right)^{n-1}\frac{1}{a}\,dx=\frac{n}{a^n}\int_0^a x^{n+1}dx=\frac{n}{n+2}a^2$$

$$\therefore\quad D\xi_n^*=\frac{n}{n+2}a^2-\left(\frac{n}{n+1}a\right)^2$$

$$=\left[\frac{n}{n+2}-\left(\frac{n}{n+1}\right)^2\right]a^2=\frac{na^2}{(n+1)^2(n+2)}$$

32　∵　(ξ,η)在三角区域△里是均匀分布

∴　$p(x,y)=\dfrac{1}{\frac{1}{2}[(3+3)\cdot 1]}=\dfrac{1}{3}$　　当$(x,y)\in$△时，这样，

$$p_1(x) = \int_{-\infty}^{\infty} p(x,y)dy$$

$$= \begin{cases} \int_{-\frac{1}{3}x}^{1} \frac{1}{3}\, dy & -3 \leqslant x \leqslant 0 \\ \int_{\frac{1}{3}x}^{1} \frac{1}{3} dy & 0 < x \leqslant 3 \end{cases}$$

$$= \begin{cases} \frac{1}{3}\left(1 + \frac{x}{3}\right) & -3 \leqslant x \leqslant 0 \\ \frac{1}{3}\left(1 - \frac{x}{3}\right) & 0 < x \leqslant 3 \end{cases}$$

$$p_2(y) = \int_{-\infty}^{\infty} p(x,y)dx = \int_{-3y}^{3y} \frac{1}{3}\, dx = \frac{1}{3}x \Big|_{-3y}^{3y}$$

$$= \frac{1}{3} \cdot 6y = 2y, \qquad 0 < y < 1,$$

$$E\xi = \int_{-\infty}^{\infty} x \cdot p_1(x)\, dx = \int_{-3}^{0} x \cdot \frac{1}{3}\left(1 + \frac{1}{3}x\right)dx +$$

$$+ \int_{0}^{3} x \cdot \frac{1}{3}\left(1 - \frac{1}{3}x\right)dx$$

$$= \frac{1}{3}\left[\frac{x^2}{2} + \frac{x^3}{9}\right]_{-3}^{0} + \frac{1}{3}\left[\frac{x^2}{2} - \frac{x^3}{9}\right]_{0}^{3}$$

$$= \frac{1}{3}\left[-\frac{9}{2} + 3\right] + \frac{1}{3}\left[\frac{9}{2} - 3\right]$$

$$= 0$$

$$E\xi^2 = \int_{-\infty}^{\infty} x^2 \cdot p_1(x)\, dx = \int_{-3}^{0} x^2 \cdot \frac{1}{3}\left(1 + \frac{1}{3}x\right)dx +$$

$$+ \int_{0}^{3} x^2 \cdot \frac{1}{3}\left(1 - \frac{1}{3}x\right)dx$$

$$= \frac{1}{3}\left[\int_{-3}^{0}\left(x^2 + \frac{1}{3}x^3\right) dx + \int_{0}^{3}\left(x^2 - \frac{1}{3}x^3\right)dx.\right.$$

$$= \frac{1}{3}\left[\left(\frac{x^3}{3} + \frac{x^4}{12}\right)\Big|_{-3}^{0} + \left(\frac{x^3}{3} - \frac{x^4}{12}\right)\Big|_{0}^{3}\right]$$

$$= \frac{1}{3}\left[\left(9 - \frac{27}{4}\right) + \left(9 - \frac{27}{4}\right)\right]$$

$$= \frac{2}{3}\left(9 - \frac{27}{4}\right)$$

$$= \frac{2}{3} \cdot \frac{36 - 27}{4}$$

$$= \frac{2}{3} \cdot \frac{9}{4}$$

$$= \frac{3}{2}$$

$$D\xi = E\xi^2 - (E\xi)^2 = \frac{3}{2} - 0^2 = 1.5$$

同理 $E\eta = \int_{-\infty}^{\infty} y p_2(y)\,dy = \int_0^1 2y^2 dy = \frac{2}{3} y^3 \Big|_0^1 = \frac{2}{3}$

$$E\eta^2 = \int_{-\infty}^{\infty} y^2 \cdot p_2(y)\,dy = 2\int_0^1 y^3 dy = \frac{y^4}{2}\Big|_0^1 = \frac{1}{2}$$

$$\therefore \quad D\eta = E\eta^2 - (E\eta)^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$$

33 （1）$\because$ (ξ,η)在三角区域△内均匀分布.

$\therefore \quad p(x,y) = \dfrac{1}{(2\times 4)\cdot\frac{1}{2}} = \dfrac{1}{4}$ 当$(x,y)\in$△时，这样，

$$p_1(x) = \int_{-\infty}^{\infty} p(x,y)\,dy$$

$$= \int_0^{2x} \frac{1}{4}\,dy$$

$$= \frac{1}{4} \cdot 2x = \frac{x}{2} \qquad 0 < x < 2;$$

$$p_2(y) = \int_{-\infty}^{\infty} p(x,y)\,dx$$

$$= \int_{\frac{y}{2}}^{2} \frac{1}{4}\,dx = \frac{1}{4}x\Big|_{\frac{y}{2}}^{2} = \frac{1}{4}\left(2 - \frac{y}{2}\right)$$

$$= \frac{1}{2} - \frac{y}{8} \qquad 0 < y < 4.$$

$$E\xi = \int_{-\infty}^{\infty} x p_1(x) dx = \int_0^2 x \cdot \frac{x}{2} dx = \frac{1}{2} \int_0^2 x^2 dx$$

$$= \frac{1}{2} \cdot \frac{x^3}{3} \Big|_0^2 = \frac{4}{3}$$

$$E\xi^2 = \int_{-\infty}^{\infty} x^2 p_1(x) dx = \int_0^2 x^2 \cdot \frac{x}{2} dx = \frac{1}{2} \int_0^2 x^3 dx$$

$$= \frac{1}{2} \cdot \frac{x^4}{4} \Big|_0^2 = 2$$

$$\therefore \quad D\xi = E\xi^2 - (E\xi)^2 = 2 - \left(\frac{4}{3}\right)^2 = 2 - \frac{16}{9} = \frac{2}{9}$$

$$E\eta = \int_{-\infty}^{\infty} y p_2(y) dy = \int_0^4 y \left(\frac{1}{2} - \frac{y}{8}\right) dy$$

$$= \frac{1}{2} \int_0^4 \left(y - \frac{y^2}{4} \right) dy = \frac{1}{2} \left[\frac{y^2}{2} - \frac{y^3}{12} \right]_0^4$$

$$= \frac{1}{2} \left[\frac{16}{2} - \frac{4^3}{12} \right] = \frac{1}{2} \left(8 - \frac{16}{3} \right) = \frac{4}{3}$$

$$E\eta^2 = \int_{-\infty}^{\infty} y^2 p_2(y) dy = \int_0^4 y^2 \left(\frac{1}{2} - \frac{y}{8} \right) dy$$

$$= \frac{1}{2} \int_0^4 \left(y^2 - \frac{y^3}{4} \right) dy = \frac{1}{2} \left(\frac{y^3}{3} - \frac{y^4}{16} \right)_0^4$$

$$= \frac{1}{2} \left(\frac{64}{3} - \frac{4^4}{16} \right) = \frac{1}{2} \left(\frac{64}{3} - 16 \right)$$

$$= \frac{32}{3} - 8 = \frac{32 - 24}{3}$$

$$= \frac{8}{3}$$

$$\therefore \quad D\eta = E\eta^2 - (E\eta)^2 = \frac{8}{3} - \left(\frac{4}{3} \right)^2$$

$$= \frac{8}{3} - \frac{16}{9} = \frac{24 - 16}{9}$$

$$= \frac{8}{9}$$

34　（1）设　$\eta=\frac{\xi-2}{0.16}$

$$\eta\sim N(0,\ 1)$$

$$P(\xi\geqslant 2.3)=P\left(\eta\geqslant\frac{2.3-2}{0.16}\right)$$

$$=P\{\eta\geqslant 1.875\}$$

$$=1-P(\eta<1.875)$$

$$=1-\Phi(1.875)$$

$$=1-0.9696$$

$$=0.0304$$

（2）$P(1.8\leqslant\xi\leqslant 2.1)$

$$=\Phi\left(\frac{2.1-2}{0.16}\right)-\Phi\left(\frac{1.8-2}{0.16}\right)$$

$$=\Phi(0.625)-\Phi(-1.25)$$

$$=\Phi(0.625)+\Phi(1.25)-1$$

$$=\frac{0.7324+0.7357}{2}+0.8914-1$$

$$=0.73405+0.8914-1$$

$$=1.62545-1$$

$$=0.62545$$

35　设测高器的误差用 ξ 来表示.

于是有

$$P(-2.5\leqslant\xi\leqslant 2.5)=0.9$$

又设　$\eta=\frac{\xi-0}{\sigma}$.

$$\eta\sim N(0,\ 1)$$

∴　$P\left(\frac{-2.5}{\sigma}\leqslant\eta\leqslant\frac{2.5}{\sigma}\right)=0.9$

即　$\Phi\left(\frac{2.5}{\sigma}\right)-\Phi\left(\frac{-2.5}{\sigma}\right)=0.9$

$$2\Phi\left(\frac{2.5}{\sigma}\right)-1=0.9$$

$$\Phi\left(\frac{2.5}{\sigma}\right)=0.95$$

反查表 $\frac{2.5}{\sigma}=1.645$

$\therefore\quad \sigma\doteq 1.52$

36 设ξ表示直径的长度.（$\xi>0$）

又设 $\eta=\frac{\xi-0.8}{0.02}$

$\eta\sim N(0,\ 1)$

$$\begin{aligned}P(\xi>0.81)&=1-P(\xi\leqslant 0.81)\\&=1-P\left(\eta\leqslant\frac{0.81-0.80}{0.02}\right)\\&=1-P(\eta\leqslant 0.5)\\&=1-\Phi(0.5)\\&=1-0.6915\\&=0.3085\end{aligned}$$

37 $\because\quad \xi\sim N(0,\ 1)$

$\therefore\quad P(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$

$$\begin{aligned}G(y)&=P(\eta\leqslant y)=P(|\xi|\leqslant y)\\&=P(-y\leqslant\xi\leqslant y)\\&=\int_{-y}^{y}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx\\&=\frac{2}{\sqrt{2\pi}}\int_0^y e^{-\frac{x^2}{2}}dx\end{aligned}$$

$\therefore\quad g(y)=G'(y)=\frac{2}{\sqrt{2\pi}}e^{-\frac{y^2}{2}}=\sqrt{\frac{2}{\pi}}e^{-\frac{y^2}{2}},\qquad y\geqslant 0$

$$E\eta=\int_0^{\infty} y\sqrt{\frac{2}{\pi}}e^{-\frac{y^2}{2}}dy=\sqrt{\frac{2}{\pi}}\left(-e^{-\frac{y^2}{2}}\right)\Big|_0^{\infty}$$

$$= 0 + \sqrt{\frac{2}{\pi}} = \sqrt{\frac{2}{\pi}}$$

$$E\eta^2 = \int_0^\infty y^2 \sqrt{\frac{2}{\pi}} e^{-\frac{y^2}{2}} dy = -\sqrt{\frac{2}{\pi}} y e^{-\frac{y^2}{2}} \Big|_0^\infty$$

$$+ \sqrt{\frac{2}{\pi}} \int_0^\infty e^{-\frac{y^2}{2}} dy = 0 + \sqrt{\frac{2}{\pi}} \cdot \frac{\sqrt{\pi}}{2\sqrt{\frac{1}{2}}} = 1$$

$\therefore$ $$D\eta = E\eta^2 - (E\eta)^2 = 1 - \left(\sqrt{\frac{2}{\pi}}\right)^2 = 1 - \frac{2}{\pi}$$

$$= \frac{\pi - 2}{\pi}$$

38 $\because$ $P(\xi \leqslant c) = 2P(\xi > c)$

$P(\xi \leqslant c) + P(\xi > c) = 3P(\xi > c)$

$\therefore$ $P(\xi > c) = \frac{1}{3}$

$1 - P(\xi \leqslant c) = \frac{1}{3}$

$P(\xi \leqslant c) = \frac{2}{3}$

设 $\eta = \frac{\xi - \mu}{\sigma}$

$\eta \sim N(0, 1)$

$P\left(\eta \leqslant \frac{c - \mu}{\sigma}\right) = 0.6667$

反查表得：

$\frac{c - \mu}{\sigma} = 0.43$

$\therefore$ $c = 0.43\sigma + \mu$

39 $$E\xi = \int_{-\infty}^{\infty} x \cdot p(x) dx$$

$$= \int_{-\infty}^{\infty} x \cdot \frac{1}{2a} e^{-\frac{|x - a|}{a}} dx$$

$$= \frac{1}{2\alpha}\left[\int_{-\infty}^{a} xe^{\frac{x-a}{\alpha}}dx + \int_{a}^{\infty} xe^{\frac{a-x}{\alpha}}dx\right]$$

$$= \frac{1}{2\alpha}\left[x \cdot \alpha e^{\frac{x-a}{\alpha}}\Big|_{-\infty}^{a} - \int_{-\infty}^{a} \alpha e^{\frac{x-a}{\alpha}}dx\right.$$

$$\left. + x\left(-\alpha e^{\frac{a-x}{\alpha}}\right)\Big|_{a}^{\infty} + \int_{a}^{\infty} \alpha e^{\frac{a-x}{\alpha}}dx\right]$$

$$= \frac{1}{2\alpha}\left[a\alpha - \left(\alpha^2 e^{\frac{x-a}{\alpha}}\right)\Big|_{-\infty}^{a} + a\alpha + \left(-\alpha^2 e^{\frac{a-x}{\alpha}}\right)\Big|_{a}^{\infty}\right]$$

$$= \frac{1}{2\alpha}(2a\alpha)$$

$$= a$$

$$E\xi^2 = \int_{-\infty}^{\infty} x^2 p(x)dx = \int_{-\infty}^{\infty} x^2 \frac{1}{2\alpha} e^{-\frac{|x-a|}{\alpha}}dx$$

$$= \frac{1}{2\alpha}\left[\int_{-\infty}^{a} x^2 e^{\frac{x-a}{\alpha}}dx + \int_{a}^{\infty} x^2 e^{\frac{a-x}{\alpha}}dx\right]$$

$$= \frac{1}{2\alpha}\left[x^2 \alpha e^{\frac{x-a}{\alpha}}\Big|_{-\infty}^{a} - \int_{-\infty}^{a} 2x\alpha e^{\frac{x-a}{\alpha}}dx\right.$$

$$\left. + \left(-x^2 \alpha e^{\frac{a-x}{\alpha}}\Big|_{a}^{\infty} + \int_{a}^{\infty} 2\alpha x e^{\frac{a-x}{\alpha}}dx\right)\right]$$

$$= \frac{1}{2\alpha}\left[a - 2\alpha\int_{-\infty}^{a} xe^{\frac{x-a}{\alpha}}dx + a\right.$$

$$\left. + 2\alpha\int_{a}^{\infty} xe^{-\frac{x-a}{\alpha}}dx\right]$$

$$= \frac{1}{2\alpha}[(a^2\alpha - 2\alpha(a\alpha - \alpha^2)) + (a^2\alpha + 2\alpha(a\alpha + \alpha^2))]$$

$$= \frac{1}{2\alpha}[a^2\alpha - 2\alpha^2 a + 2\alpha^3 + a^2\alpha + 2\alpha^2 a + 2\alpha^3]$$

$$= \frac{1}{2\alpha}[2a^2\alpha + 4\alpha^3]$$

$$= a^2 + 2\alpha^2$$

$$\therefore \quad D\xi = E\xi^2 - (E\xi)^2 = (a^2 + 2\alpha^2) - a^2$$

$$= 2\alpha^2$$

40 $P(1<\xi^2<4)$

$= P(1<|\xi|<2)$

$= P([(\xi<-1)\cup(\xi>1)]\cap[-2<\xi<2])$

$= P((-2<\xi<-1)\cup(1<\xi<2))$

$= P\{-2<\xi<-1\}+P\{1<\xi<2\}$

设 $\eta=\dfrac{\xi-0}{5}$.

$\eta\sim N(0,\ 1)$.

$\therefore$ 原式 $=P(-0.4<\eta<-0.2)+P(0.2<\eta<0.4)$

$=\Phi(-0.2)-\Phi(-0.4)+\Phi(0.4)-\Phi(0.2)$

$=1-\Phi(0.2)-1+\Phi(0.4)+\Phi(0.4)-\Phi(0.2)$

$=2[\Phi(0.4)-\Phi(0.2)]$

$=2(0.6554-0.5793)$

$=2\times0.0761$

$=0.1522$

41 $\because$ ξ 服从参数为a的右截尾指数分布。

$$\therefore\quad p(x)=\begin{cases}cae^{-ax} & 0<x<T\\ 0 & \text{其它}\end{cases}$$

其中$a>0$，$c=[1-e^{-aT}]^{-1}$.

因此，$E\xi=c\displaystyle\int_0^T x\cdot ae^{-ax}dx$

$$=c\left[\left(-xe^{-ax}\right)\Big|_0^T+\int_0^T e^{-ax}dx\right]$$

$$=c\left[-Te^{-aT}-\frac{1}{a}\int_0^T de^{-ax}\right]$$

$$=c\left[-Te^{-aT}-\frac{1}{a}e^{-aT}+\frac{1}{a}\right]$$

$$=\frac{1}{1-e^{-aT}}\left[\frac{1}{a}-e^{-aT}\left(T+\frac{1}{a}\right)\right]$$

42 令$N(0,1)$的密度函数为$f(x)$，则(ξ_1,ξ_2)的联合分布密度

$$f(x_1,x_2)=f(x_1)f(x_2)=\frac{1}{2\pi}e^{-\frac{x_1^2+x_2^2}{2}},\quad x_1,x_2\in(-\infty,\infty)$$

令 $z=x_1+x_2$，$w=x_1-x_2$，则$x_1=\frac{z+w}{2}$，$x_2=\frac{z-w}{2}$.

由定理3.3.3，二维随机变量(Z,W)的联合分布密度

$$g(z,w)=\frac{1}{2\pi}e^{-\frac{x_1^2+x_2^2}{2}}|J|$$

其中 $J=\begin{vmatrix}\frac{1}{2} & \frac{1}{2}\\ \frac{1}{2} & -\frac{1}{2}\end{vmatrix}=-\frac{1}{2}$,而 $x_1^2+x_2^2=\frac{z^2+2zw+w^2}{4}+\frac{z^2-2zw+w^2}{4}=\frac{z^2+w^2}{2}$

所以

$$g(z,w)=\frac{1}{4\pi}e^{-(z^2+w^2)/4}\qquad z,w\in(-\infty,\infty)$$

Z的边缘分布密度

$$g_1(z)=\int_{-\infty}^{\infty}\frac{1}{4\pi}e^{-(z^2+w^2)/4}dw=\frac{1}{2\sqrt{\pi}}e^{-\frac{z^2}{4}}$$

W的边缘分布密度

$$g_2(w)=\int_{-\infty}^{\infty}\frac{1}{4\pi}e^{-(z^2+w^2)/4}dz=\frac{1}{2\sqrt{\pi}}e^{-\frac{w^2}{4}}$$

于是 $g(z,w)=g_1(z)\cdot g_2(w)$,所以得证$Z$与$W$相互独立。

43 因 $\xi_1-\bar{\xi}=\xi_1-\frac{\xi_1+\xi_2}{2}=\frac{\xi_1-\xi_2}{2}$

$$\xi_2-\overline{\xi}=\frac{\xi_2-\xi_1}{2}=-\frac{\xi_1-\xi_2}{2}$$

所以 $$S^2=\frac{(\xi_1-\xi_2)^2}{2}$$

运用42题的符号，则

$$\overline{\xi}=\frac{Z}{2},\qquad S^2=\frac{W^2}{2}$$

由42题知Z，W相互独立，所以$\overline{\xi}$与S^2分别作为Z和W的函数也是相互独立的。于是由Z和W的密度函数立刻推出$\overline{\xi}$和S^2的分布密度及其联合分布密度：

$\overline{\xi}$的分布密度 $h_1(x)=\dfrac{1}{\sqrt{\pi}}e^{-x^2}\qquad x\in(-\infty,\infty)$

S^2的分布密度 $h_2(y)=\dfrac{1}{\sqrt{2\pi}}y^{-\frac{1}{2}}e^{-\frac{y}{2}}\qquad y\in(0,\infty)$

其联合分布密度

$$h(x,y)=\frac{1}{\pi\sqrt{2}}e^{-x^2}y^{-\frac{1}{2}}e^{-\frac{y}{2}}\qquad -\infty<x<\infty,0<y<\infty$$

由$h_1(x)$和$h_2(y)$的表达式可以看出，$\overline{\xi}\sqrt{2}\sim N(0,\ 1)$，$S_2\sim\chi^2(1)$，根据定理3.5.10，随机变量 $\zeta=\overline{\xi}\sqrt{2}/\sqrt{S^2}$服从自由度为1的$t$分布，其分布密度为

$$p(t)=\frac{1}{\pi}\cdot\frac{1}{1+t^2}\qquad t\in(-\infty,\ \infty)$$

第四章　极 限 定 理

在实际中，常常需要考虑许多随机因素的总的影响。比如我们研究炮弹弹着点的分布问题，一枚炮弹究竟落到哪，那是由发射条件及大量随机因素决定的，像瞄准时的误差 ξ_1，炮弹装药量的误差 ξ_2，炮身结构引起的误差 ξ_3，空气阻力产生的误差 ξ_4， 等等都对炮弹的弹着点有影响，而这些随机变量的总和 $\Sigma\xi_i$ 就决定了炮弹弹着点位置的误差。因此研究随机变量和的极限行为是一个很重要的课题，这就是本章的中心内容。

§4.1 依分布收敛与依概率收敛

研究随机变量的极限行为，首先碰到的问题就是在什么意义下收敛的问题。在这一节里，我们考虑两类收敛性的模型——依分布收敛和依概率收敛，给出这两种收敛性的定义并讨论它们之间的关系。

定义 4.1.1　设 $\{F_n\}$ 是分布函数序列，假如存在单调不减函数 F，使得当 $n\to\infty$ 时，$F_n(x)\to F(x)$ 对 F 的每一个连续点 x 成立，我们说 F_n 弱收敛到 F，记为 $F_n\xrightarrow{W}F$。

若 $\{\xi_n\}$ 是随机变量序列，$\{F_n\}$ 是其对应的分布函数序列。如果存在一个随机变量 ξ，使得 F_n 弱收敛到 ξ 的分布函数 F，则称 ξ_n 依分布收敛到 ξ，记为 $\xi_n\xrightarrow{L}\xi$。

注意，在弱收敛定义里没有要求 F_n 一定收敛到一个分布函数

F，给定一个分布函数序列，它完全有可能不收敛到一个分布函数。

例4.1.1 考虑分布函数序列

$$F_n(x)=\begin{cases}0 & x<n\\ 1 & x\geqslant n\end{cases}$$

这里 $F_n(x)$是只取值于单点 n 的退化随机变量的分布函数．它收敛于恒为 0 的函数 F，显然 F 不是分布函数．

例 4.1.2 设 $\xi_1, \xi_2, \cdots, \xi_n$ 独立同分布，其共同的分布密度为

$$f(x)=\begin{cases}\dfrac{1}{\theta} & 0<x<\theta\\ 0 & 其\quad 它\end{cases}\qquad(0<\theta<\infty)$$

令 $\xi_n^*=\max(\xi_1, \xi_2, \cdots, \xi_n)$，则$\xi_n^*$的密度函数为

$$f_n(x)=\begin{cases}\dfrac{nx^{n-1}}{\theta^n} & 0<x<\theta\\ 0 & 其\quad 它\end{cases}$$

ξ_n^*的分布函数为（见定理3.3.5）

$$F_n(x)=\begin{cases}0 & x<0\\ (x/\theta)^n & 0\leqslant x<\theta\\ 1 & x\geqslant\theta\end{cases}$$

当 $n\to\infty$时，$F_n(x)\to F(x)=\begin{cases}0 & x<\theta\\ 1 & x\geqslant\theta\end{cases}$

$F(x)$ 是分布函数．

下面的例子指出依分布收敛并不意味着对应的分布列或分布密度也收敛．

例4.1.3 令$\{\xi_n\}$的分布列为

$$f_n(x)=P(\xi_n=x)=\begin{cases}1 & x=2+\dfrac{1}{n}\\ 0 & 其\quad 它\end{cases}$$

注意，对任何一个 $f_n(x)$，当 $x=2$ 时均恒为 0，

$$f_n(x)\to f(x)\quad(n\to\infty)$$

这时对一切 x，$f(x)=0$。然而$\{F_n\}$弱收敛到分布函数

$$F(x)=\begin{cases}0 & x<2\\ 1 & x\geqslant 2\end{cases}$$

容易证明下面定理。

定理4.1.1　令ξ_n是只取整数值的随机变量，$f_n(k)=P(\xi_n=k)$，$k=0,1,2,\cdots$，是ξ_n的分布列，$n=1,2,\cdots$，$f(k)=P(\xi=k)$是随机变量ξ的分布列，则

$$f_n(x)\to f(x) \text{ 对一切 } x \text{ 成立} \Longleftrightarrow \xi_n\xrightarrow{L}\xi.$$

其中 $f_n(x)$和 $f(x)$当 $x\neq k(k=0,1,2\cdots)$时均为零。

对连续情形我们有下面结果

定理4.1.2　若ξ_n，$n=1,2,\cdots$，和ξ是连续型随机变量，且对（几乎）所有的x，有

$$f_n(x)\to f(x)\quad (n\to\infty)$$

这里 f_n 和 f 分别是ξ_n 和ξ的分布密度，则 $\xi_n\xrightarrow{L}\xi$。

定理4.1.3　若$\{\xi_n\}$是随机变量序列，$\xi_n\xrightarrow{L}\xi$，c是常数，则

（a）$\xi_n+c\xrightarrow{L}\xi+c$

（b）$c\xi_n\xrightarrow{L}c\xi\,(c\neq 0)$

定义4.1.2　　令$\{\xi_n\}$是随机变量序列，如果存在随机变量ξ，使得对任意 $\varepsilon>0$，当 $n\to\infty$时，有

$$P(|\xi_n-\xi|>\varepsilon)\to 0$$

则称ξ_n依概率收敛到ξ，记为$\xi_n\xrightarrow{P}\xi$。

注意，$\xi_n\xrightarrow{P}\xi$并不意味着对给定$\varepsilon>0$．可以找到N、使得当$n>N$时，$|\xi_n-\xi|<\varepsilon$，而只是说概率序列$P(|\xi_n-\xi|>\varepsilon)\to 0$。

例4.1.4　令$\{\xi_n\}$是随机变量序列，其分布列为

$$P(\xi_n=1)=\frac{1}{n},\quad P(\xi_n=0)=1-\frac{1}{n}$$

则
$$P(|\xi_n|>\varepsilon)=\begin{cases}P(\xi_n=1)=\frac{1}{n} & \text{当}0<\varepsilon<1\\ 0 & \varepsilon\geqslant 1\end{cases}$$

由此可得当 $n\to\infty$ 时，$P(|\xi_n|>\varepsilon)\to 0$ 故 $\xi_n\xrightarrow{P}0$。

下面诸结论都是容易验证的。

1　$\xi_n\xrightarrow{P}\xi\Longleftrightarrow\xi_n-\xi\xrightarrow{P}0$，

2　$\xi_n\xrightarrow{P}\xi$，$\xi_n\xrightarrow{P}\eta\Rightarrow P(\xi=\eta)=1$。

这是因为 $P(|\xi-\eta|>c)\leqslant P\left(|\xi_n-\xi|>\frac{c}{2}\right)+P\left(|\xi_n-\eta|>\frac{c}{2}\right)$

这样，对任意 $c>0$，有 $P(|\xi-\eta|>c)=0$。

3　$\xi_n\xrightarrow{P}\xi\Rightarrow\xi_n-\xi_m\xrightarrow{P}0$，当 $n,m\to\infty$ 时。这是因为

$$P(|\xi_n-\xi_m|>\varepsilon)\leqslant P\left(|\xi_n-\xi|>\frac{\varepsilon}{2}\right)+P\left(|\xi_m-\xi|>\frac{\varepsilon}{2}\right)$$

4　$\xi_n\xrightarrow{P}\xi$，$\eta_n\xrightarrow{P}\eta\Rightarrow\xi_n\pm\eta_n\xrightarrow{P}\xi\pm\eta$。

5　$\xi_n\xrightarrow{P}\xi$，k 为常数$\Rightarrow k\xi_n\xrightarrow{P}k\xi$。

6　$\xi_n\xrightarrow{P}\xi\Rightarrow\xi_n^2\xrightarrow{P}\xi^2$。

7　$\xi_n\xrightarrow{P}a$，$\eta_n\xrightarrow{P}b$，a,b 为常数 $\Rightarrow\xi_n\eta_n\xrightarrow{P}ab$。

原因是

$$\xi_n\eta_n=\frac{(\xi_n+\eta_n)^2-(\xi_n-\eta_n)^2}{4}\xrightarrow{P}\frac{(a+b)^2-(a-b)^2}{4}$$
$$=ab.$$

8　$\xi_n\xrightarrow{P}1\Rightarrow\xi_n^{-1}\xrightarrow{P}1$。

原因是

$$P\left(\left|\frac{1}{\xi_n}-1\right|\geqslant\varepsilon\right)=P\left(\frac{1}{\xi_n}\geqslant 1+\varepsilon\right)+P\left(\frac{1}{\xi_n}\leqslant 1-\varepsilon\right)$$

$$= P\left(\frac{1}{\xi_n} \geqslant 1+\varepsilon\right) + P\left(\frac{1}{\xi_n} \leqslant 0\right)$$

$$+ P\left(0 < \frac{1}{\xi_n} \leqslant 1-\varepsilon\right)$$

右边三项中每一个当 $n\to\infty$ 时都趋于 0 .

9 $\xi_n \xrightarrow{P} a$，$\eta_n \xrightarrow{P} b$，(a,b为常数，$b\neq 0$)$\Rightarrow \xi_n\eta_n^{-1} \xrightarrow{P} ab^{-1}$.

10 $\xi_n \xrightarrow{P} \xi$，$\eta$ 是随机变量$\Rightarrow \xi_n\eta \xrightarrow{P} \xi\eta$.

因 η 是随机变量，所以对任意给定的 $\delta>0$，存在 $k>0$，使得 $P(|\eta|>k)<\frac{\delta}{2}$，于是

$$\begin{aligned}P(|\xi_n\eta-\xi\eta|>\varepsilon) &= P(|\xi_n-\xi||\eta|>\varepsilon,\ |\eta|>k)\\ &\quad + P(|\xi_n-\xi||\eta|>\varepsilon,\ |\eta|<k)\\ &< \frac{\delta}{2} + P\left(|\xi_n-\xi|>\frac{\varepsilon}{k}\right)\end{aligned}$$

11 $\xi_n \xrightarrow{P} \xi$，$\eta_n \xrightarrow{P} \eta \Rightarrow \xi_n\eta_n \xrightarrow{P} \xi\eta$

因 $(\xi_n-\xi)(\eta_n-\eta) \xrightarrow{P} 0$，左边乘开，利用10即可得出要证结果.

定理 4.1.4 令$\xi_n \xrightarrow{P} \xi$，$g$ 是 R 上连续函数，则当 $n\to\infty$ 时 $g(\xi_n) \xrightarrow{P} g(\xi)$.

证明 因 ξ 是随机变量，所以给定 $\varepsilon>0$，存在 $k=k(\varepsilon)$，使得 $P(|\xi|>k)\leqslant\frac{\varepsilon}{2}$，又因 g 是连续函数，所以在$[-k,k]$上一致连续，于是存在 $\delta=\delta(\varepsilon,k)$，使得对一切$|x|\leqslant k$，$|x_n-x|<\delta$，有 $|g(x_n)-g(x)|<\varepsilon$. 令

$$A=(|\xi|\leqslant k),\ B=(|\xi_n-\xi|<\delta),\ C=(|g(\xi_n)-g(\xi)|<\varepsilon),$$

则 $\omega\in A\cap B\Rightarrow\omega\in C$，因此 $A\cap B\subset C$，于是

$$P(\overline{C}) \leqslant P(\overline{A}) + P(\overline{B})$$

这样

$$P(|g(\xi_n)-g(\xi)|\geqslant\varepsilon)\leqslant P(|\xi_n-\xi|\geqslant\delta)+P(|\xi|>k)<\varepsilon$$

对 $n\geqslant N(\varepsilon,\delta,k)$成立，其中 $N(\varepsilon,\delta,k)$选得使 $n\geqslant N(\varepsilon,\delta,k)$时

$$P(|\xi_n-\xi|\geqslant\delta)<\frac{\varepsilon}{2}$$

成立。

推论 $\xi_n\xrightarrow{P}c$，c 为常数$\Rightarrow g(\xi_n)\xrightarrow{P}g(c)$，其中 g 是连续函数。

下面两个定理给出了依分布收敛和依概率收敛间的关系。

定理4.1.5 $\xi_n\xrightarrow{P}\xi\Rightarrow\xi_n\xrightarrow{L}\xi$。

证明 令 F_n，F 分别是 ξ_n 和 ξ 的分布函数，我们有

$$\{\omega:\xi(\omega)\leqslant x'\}=\{\omega:\xi_n(\omega)\leqslant x,\xi(\omega)\leqslant x'\}\cup\{\omega:\xi_n(\omega)>x,\ \xi(\omega)\leqslant x'\}\subset\{\xi_n\leqslant x\}\cup\{\xi_n>x,\xi\leqslant x'\}$$

由此推出

$$F(x')\leqslant F_n(x)+P(\xi_n>x,\ \xi\leqslant x')$$

因 $\xi_n\xrightarrow{P}\xi$，所以对 $x'<x$，当 $n\to\infty$ 时，我们有

$$P(\xi_n>x,\ \xi\leqslant x')\leqslant P(|\xi_n-\xi|>x-x')\to 0$$

因此 $F(x')\leqslant \varliminf_{n\to\infty} F_n(x)$，$x'<x$

类似地，我们有

$$\varlimsup_{n\to\infty} F_n(x)\leqslant F(x''),\quad x<x''$$

这样，对 $x'<x<x''$，我们有

$$F(x')\leqslant \varliminf F_n(x)\leqslant \varlimsup F_n(x)\leqslant F(x'')$$

因 F 只有可数个不连续点，我们选 x 是 F 的连续点，且 $x''\downarrow x$，$x'\uparrow x$，所以

$$F(x)=\lim_{n\to\infty}F_n(x)$$

在 F 的一切连续点成立。

定理 4.1.6 令 k 是常数，则

$$\xi_n\xrightarrow{L}k\Rightarrow\xi_n\xrightarrow{P}k$$

证明留给读者作为练习。

注意，我们不能把上面结果改进成把其中常数 k 换成随机变量，即一般地 $\xi_n \xrightarrow{L} \xi$，不意味着 $\xi_n \xrightarrow{P} \xi$，比如，当 $\xi, \xi_1, \xi_2, \cdots$ 相互独立时，(ξ_n, ξ) 的联合分布如下

ξ \ ξ_n	0	1	
0	0	$\frac{1}{2}$	$\frac{1}{2}$
1	$\frac{1}{2}$	0	$\frac{1}{2}$
	$\frac{1}{2}$	$\frac{1}{2}$	1

显然 $\xi_n \xrightarrow{L} \xi$，但

$$P\left(|\xi_n-\xi|>\frac{1}{2}\right) \geqslant P(|\xi_n-\xi|=1)$$
$$=P(\xi_n=0,\ \xi=1)+P(\xi_n=1,\ \xi=0)$$
$$=1$$

因此 ξ_n 不依概率 P 收敛于 ξ，

§4.2 矩母函数

矩母函数是讨论随机变量和的一个有力的工具，同时它也给出了随机变量的分布与其各阶矩间的关系。

对每一个随机变量我们都可以根据其分布来定义它的矩母函数。下面分别对离散型随机变量和连续型随机变量给出其矩母函数的定义。

定义4.2.1 （1）设 ξ 是离散型随机变量，其分布列为

$$\begin{pmatrix} x_1, & x_2, & \cdots & x_n, & \cdots \\ p(x_1), & p(x_2), & \cdots & p(x_n), & \cdots \end{pmatrix}$$

则函数

$$M_\xi(t)=\sum_{i=1}^{\infty}e^{tx_i}p(x_i)$$

称为ξ的矩母函数。

（2）若ξ是具有分布密度为 $f(x)$ 的连续型随机变量，则定义其矩母函数为

$$M_\xi(t)=\int_{-\infty}^{\infty}e^{tx}f(x)dx^*$$

由上面定义可以看出，对任一随机变量ξ，其矩母函数可写成

$$M_\xi(t)=E(e^{t\xi})。$$

例4.2.1 设ξ是在[a, b]上均匀分布的随机变量，则

$$M_\xi(t)=\int_a^b\frac{e^{tx}}{b-a}\,dx=\frac{1}{t(b-a)}[e^{bt}-e^{at}]$$

例4.2.2 设ξ服从参数为n， p的二项分布，则其矩母函数为

$$\begin{aligned}M_\xi(t)&=\sum_{k=0}^{n}e^{tk}\binom{n}{k}p^k(1-p)^{n-k}\\&=\sum_{k=0}^{n}\binom{n}{k}(pe^t)^k(1-p)^{n-k}\\&=[pe^t+(1-p)]^n\end{aligned}$$

例4.2.3 设ξ服从参数为λ的$Poisson$分布，则

$$\begin{aligned}M_\xi(t)&=\sum_{k=0}^{\infty}e^{tk}\frac{\lambda^ke^{-\lambda}}{k!}=e^{-\lambda}\sum_{k=0}^{\infty}\frac{(\lambda e^t)^k}{k!}\\&=e^{-\lambda}e^{\lambda e^t}=e^{\lambda(e^t-1)}\end{aligned}$$

例4.2.4 设ξ服从参数为α的指数分布，则

$$M_\xi(t)=\int_0^{\infty}e^{tx}\alpha e^{-\alpha x}dx=\alpha\int_0^{\infty}e^{(t-\alpha)x}dx$$

上面积分只有当 $t<\alpha$ 时收敛，所以只有这样的t值矩母函数才存在，于是

---

* 由矩母函数定义可以看出，不是对任一随机变量其矩母函数对一切t都存在。但对任一随机变量其矩母函数在$t=0$时都存在，且都等于1。

$$M_\xi(t)=\frac{a}{t-a}e^{x(t-a)}\Big|_0^\infty=\frac{a}{a-t}\qquad 当\ t<a\ 时。$$

例4.2.5　设 $\xi\sim N(\mu,\sigma^2)$

则

$$M_\xi(t)=\frac{1}{\sigma\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{tx}exp\left(-\frac{1}{2}\left[\frac{x-\mu}{\sigma}\right]^2\right)dx$$

令 $(x-\mu)/\sigma=s$，这时 $x=\sigma s+\mu$，$dx=\sigma ds$，因此

$$M_\xi(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}exp[t(\sigma s+\mu)]e^{-s^2/2}ds$$

$$=e^{t\mu}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}exp\left\{-\frac{1}{2}(s^2-2\sigma ts)\right\}ds$$

$$=e^{t\mu}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}exp\left\{-\frac{1}{2}[(s-\sigma t)^2-\sigma^2t^2]\right\}ds$$

$$=e^{t\mu+\sigma^2t^2/2}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}exp\left(-\frac{1}{2}[s-\sigma t]^2\right)ds$$

令 $s-\sigma t=v$，于是 $ds=dv$，则

$$M_\xi(t)=e^{t\mu+\sigma^2t^2/2}\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty}e^{-\frac{v^2}{2}}dv$$

$$=e^{t\mu+\sigma^2t^2/2}$$

例4.2.6 设 ξ 服从参数为 a 和 γ 的加马分布，则

$$M_\xi(t)=\frac{a}{\Gamma(\gamma)}\int_0^{\infty}e^{tx}(ax)^{\gamma-1}e^{-ax}dx$$

$$=\frac{a^\gamma}{\Gamma(\gamma)}\int_0^{\infty}x^{\gamma-1}e^{-x(a-t)}dx$$

这个积分当 $a>t$ 时收敛，令 $x(a-t)=u$，这样$dx=du/(a-t)$，则

$$M_\xi(t)=\frac{a^\gamma}{(a-t)\Gamma(\gamma)}\int_0^{\infty}\left(\frac{u}{a-t}\right)^{\gamma-1}e^{-u}du$$

$$=\left(\frac{a}{a-t}\right)^\gamma\frac{1}{\Gamma(\gamma)}\int_0^{\infty}u^{\gamma-1}e^{-u}du$$

$$=\left(\frac{a}{a-t}\right)^\gamma$$

现在我们指出为什么称 $M_\xi(t)$为矩母函数。把 e^{tx} 展成马克劳林级数

$$e^{tx}=1+tx+\frac{(tx)^2}{2!}+\cdots+\frac{(tx)^n}{n!}+\cdots$$

于是

$$M_\xi(t)=E(e^{t\xi})=E\left(1+t\xi+\frac{(t\xi)^2}{2!}+\cdots+\frac{(t\xi)^n}{n!}+\cdots\right)$$

$$=1+tE(\xi)+\frac{t^2}{2!}E(\xi^2)+\cdots+\frac{t^nE(\xi^n)}{n!}+\cdots$$

$$M'_\xi(t)=E(\xi)+tE(\xi^2)+\frac{t^2E(\xi^3)}{2!}+\cdots+\frac{t^{n-1}E(\xi^n)}{(n-1)!}+\cdots$$

令 $t=0$,

$$M'_\xi(0)=E(\xi)。$$

$$M''_\xi(t)=E(\xi^2)+tE(\xi^3)+\cdots+\frac{t^{n-2}E(\xi^n)}{(n-2)!}+\cdots$$

令 $t=0$, $M''_\xi(0)=E(\xi^2)$。

继续使用上面方法，我们可推出下述定理

定理4.2.1　若随机变量 ξ 的矩母函数 $M_\xi(t)$对于$t\in(-t_0,t_0)$存在($t_0>0$)，则有

$$M_\xi^{(n)}(t)\big|_{t=0}=E\xi^n$$

由此可见，随机变量的各阶矩可看成由函数 $M_\xi(t)$产生的，故称之为矩母函数。

例4.2.7 设 ξ 服从参数为 n, p 的二项分布，这时

$$M_\xi(t)=[pe^t+q]^n$$

$$M'_\xi(t)=n[pe^t+q]^{n-1}pe^t$$

$$M''_\xi(t)=np[e^t(n-1)(pe^t+q)^{n-2}pe^t+(pe^t+q)^{n-1}e^t]$$

因此　$E(\xi)=M'_\xi(0)=np$

$$E(\xi^2)=M''_\xi(0)=np[(n-1)p+1]$$

$$D\xi=M''_\xi(0)-[M'_\xi(0)]^2=np(1-p)$$

例4.2.8 $\xi\sim N(\alpha,\beta^2)$

$$M_\xi(t)=exp\left(\alpha t+\frac{1}{2}\beta^2t^2\right)$$

$$M'_\xi(t)=e^{\alpha t+\beta^2t^2/2}(\beta^2t+\alpha)$$

$$M''_\xi(t)=e^{\beta^2t^2/2+\alpha t}\beta^2+(\beta^2t+\alpha)^2e^{\beta^2t^2/2+\alpha t}$$

于是 $M'(0)=\alpha$，$M''(0)=\beta^2+\alpha^2$，得出 $E(\xi)=\alpha$，$D\xi=\beta^2$。

下面利用矩母函数的方法来求几何分布随机变量的均值和方差。

例4.2.9 设ξ服从几何分布，$p(\xi=k)=q^{k-1}p$，$k=1,2,\cdots$，$(p+q=1)$。这时

$$M_\xi(t)=\sum_{k=1}^{\infty}e^{tk}q^{k-1}p=\frac{p}{q}\sum_{k=1}^{\infty}(qe^t)^k$$

设 t 满足 $0<qe^t<1$，即 $t<ln(1/q)$ 则上面级数可做为几何级数求和，

$$M_\xi(t)=\frac{p}{q}qe^t[1+qe^t+(qe^t)^2+\cdots]$$

$$=\frac{p}{q}\cdot\frac{qe^t}{1-qe^t}=\frac{pe^t}{1-qe^t}$$

$$M'_\xi(t)=\frac{(1-qe^t)pe^t-pe^t(-qe^t)}{(1-qe^t)^2}=\frac{pe^t}{(1-qe^t)^2}$$

$$M''_\xi(t)=\frac{(1-qe^t)^2pe^t-pe^t2(1-qe^t)(-qe^t)}{(1-qe^t)^4}$$

$$=\frac{pe^t(1+qe^t)}{(1-qe^t)^3}$$

所以 $$E(\xi)=M'(0)=p/(1-q)^2=\frac{1}{p}$$

$$E(\xi^2)=M''(0)=p(1+q)/(1-q)^3=(1+q)/p^2$$

$$D\xi=(1+q)/p^2-(1/p)^2=\frac{q}{p^2}$$

下面的几个定理在矩母函数的应用里是特别重要的。

定理 4.2.2 设随机变量ξ的矩母函数为$M_\xi(t)$，令 $\eta=\alpha\xi+\beta$，则η的矩母函数 $M_\eta(t)$为

$$M_\eta(t)=e^{\beta t}M_\xi(\alpha t)$$

证明 $$M_\eta(t)=E(e^{t\eta})=E(e^{(\alpha\xi+\beta)t})$$
$$=e^{\beta t}E(e^{\alpha\xi t})=e^{\beta t}M_\xi(\alpha t)$$

定理4.2.3 （唯一性定理）设随机变量ξ，η的矩母函数分别为$M_\xi(t)$和$M_\eta(t)$。若对一切t值有$M_\xi(t)=M_\eta(t)$，则ξ与η具有相同的概率分布。

证明略去*. 这个定理告诉我们若两个随机变量的矩母函数相同，则它们就有相同的概率分布，就是说，矩母函数唯一确定随机变量的概率分布。

例4.2.10 设$\xi\sim N(\mu,\ \sigma^2)$，$\eta=\alpha\xi+\beta$，由第三章已知$\eta$也服从正态分布，这个事实也可以通过求$\eta$的矩母函数看出，由例4.2.5知，$\xi$的矩母函数为

$$M_\xi(t)=e^{\mu t+\sigma^2t^2/2}$$

而 $$M_\eta(t)=e^{\beta t}[e^{\alpha\mu t+(\alpha\sigma)^2t^2/2}]$$
$$=e^{(\beta+\alpha\mu)t}e^{(\alpha\sigma)^2t^2/2}$$

这正是均值为$\beta+\alpha\mu$，方差为$\alpha^2\sigma^2$的正态分布的矩母函数，由定理4.2.3就推出η服从$N(\beta+\alpha\mu,\ \alpha^2\sigma^2)$。

定理4.2.4 设ξ和η是两个相互独立的随机变量，$\zeta=\xi+\eta$，令$M_\xi(t)$，$M_\eta(t)$和$M_\zeta(t)$分别为ξ，η和ζ的矩母函数，则$M_\zeta(t)=M_\xi(t)\cdot M_\eta(t)$。

证明 $$M_\zeta(t)=E(e^{\zeta t})=E[e^{(\xi+\eta)t}]=E(e^{\xi t}e^{\eta t})$$
$$=E(e^{\xi t})\cdot E(e^{\eta t})=M_\xi(t)M_\eta(t).$$

注 这个定理可进一步推广：若$\xi_1,\xi_2,\cdots,\xi_n$是n个相互独立随机变量，其矩母函数分别是$M_{\xi_1}(t),\cdots,M_{\xi_n}(t)$，又设$M_\zeta(t)$是$\zeta=\xi_1+\xi_2+\cdots+\xi_n$的矩母函数，则$M_\zeta(t)=M_{\xi_1}(t)\cdots M_{\xi_n}(t)$。

定理4.2.5 （连续性定理）设$\{F_n\}$是分布函数序列，其对应

---

*) 此定理证明可参看 D. V. Widder, Advanced Calculus 2nd Ed., Prentice-Hall Englewood cliffs, N.J., 1961. P460.

的矩母函数序列为 $\{M_n\}$。又设对每个 n 及 $|t|\leqslant t_0$ 时 $M_n(t)$ 存在，假如存在一个分布函数 F，其对应的矩母函数 M 对于 $|t|\leqslant t_1<t_0$ 的 t 存在，并使得当 $n\to\infty$ 时 $M_n(t)\to M(t)$ 对每个 $t\in[-t_1,t_1]$ 成立，则 $F_n\xrightarrow{W}F$。

证明从略*.

定理 4.2.5 说明，为了求极限分布，可研究对应的矩母函数的极限，由矩母函数的唯一性推知此矩母函数的极限对应的分布即为所求。

§4.3 中心极限定理

有了矩母函数这一预备知识，下面我们就讨论本章开始时提出的课题——研究随机变量和的极限分布问题。设 $\xi_1,\xi_2,\cdots,\xi_n,\cdots$ 是一随机变量序列，令 $\zeta_n=\sum_{i=1}^{n}\xi_i$，则研究当 $n\to\infty$ 时，ζ_n 的渐近分布的一切命题一般都统称极限定理，而本节所要讨论的中心极限定理，是其中最重要的一种极限定理。

那么什么是中心极限定理呢？中心极限定理就是研究随机变量和渐近正态分布的条件的一系列定理的统称。从第三章的学习中我们知道正态分布是一种极其重要的分布，而在实际中大量随机现象是服从正态分布的，那么为什么会出现这种现象呢？中心极限定理就具体地从理论上说明了这个问题。

中心极限定理包括的定理很多，这里我们只介绍几个简单但十分重要的定理。

定理 4.3.1 （同分布中心极限定理）设 $\xi_1,\xi_2,\cdots$是相互独立同分布随机变量序列，又设 $E\xi_k=\mu$， $D\xi_k=\sigma^2$ 均有限。令 $\zeta_n=$

---

*) 参看E. Lukacs, *Characteristic Functions*. 2nd Ed., Hafner New York 1970, p 47.

$\sum_{i=1}^{n}\xi_i$， $T_n=\frac{\zeta_n-E\zeta_n}{\sqrt{D\zeta_n}}$，则

当 $n\to\infty$时有

$$T_n\xrightarrow{L}Z$$

其中Z是 $N(0,1)$ 随机变量.

证明　设 T_n 的分布函数为 F_n， $N(0,1)$随机变量的分布函数为Φ，为证定理我们只要证 $F_n\xrightarrow{W}\Phi$即可. 下面仅就ξ_i 的矩母函数存在的情形给出定理的证明.

令$M(t)$是ξ_k 的共同的矩母函数，因 ξ_k 相互独立，所以 ζ_n 的矩母函数 $M_{\zeta_n}(t)=[M(t)]^n$，又由于$T_n=\frac{\zeta_n-E\zeta_n}{\sqrt{D\zeta_n}}=\frac{\zeta_n-n\mu}{\sqrt{n}\sigma}$是 ζ_n 的线性函数，由定理 4.2.2 知 T_n 的矩母函数

$$M_{T_n}(t)=e^{-(\sqrt{n}\mu/\sigma)t}\left[M\left(\frac{t}{\sqrt{n}\sigma}\right)\right]^n$$

于是
$$\ln M_{T_n}(t)=\frac{-\sqrt{n}\mu}{\sigma}t+n\ln M\left(\frac{t}{\sqrt{n}\sigma}\right)$$

把 $M(t)$展成马克劳林级数，

$$M(t)=1+M'(0)t+\frac{M''(0)t^2}{2}+R$$

其中R是余项。因 $M'(0)=\mu,M''(0)=\mu^2+\sigma^2$，故

$$M(t)=1+\mu t+\frac{(\mu^2+\sigma^2)t^2}{2}+R$$

所以

$$\ln M_{T_n}(t)=-\frac{\sqrt{n}\mu t}{\sigma}+n\ln\left[1+\frac{\mu t}{\sqrt{n}\sigma}+\frac{(\mu^2+\sigma^2)t^2}{2n\sigma^2}+R\right]$$

我们利用 $\ln(1+x)$的马克劳林展开式

$$\ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}+\cdots$$

（这个展开式当 $|x|<1$ 是正确的，在我们的问题里，$x=\frac{\mu t}{\sqrt{n}\sigma}+\frac{(\mu^2+\sigma^2)t^2}{2n\sigma^2}+R$，而对充分大的 n，这个表达式的绝对值将小于 1）这样我们有

$$\ln M_{T_n}(t)=-\frac{\sqrt{n}\mu t}{\sigma}+n\left[\left(\frac{\mu t}{\sqrt{n}\sigma}+(\mu^2+\sigma^2)\frac{t^2}{2n\sigma^2}+R\right)-\frac{1}{2}\left(\frac{\mu t}{\sqrt{n}\sigma}+(\mu^2+\sigma^2)\frac{t^2}{2n\sigma^2}+R\right)^2+\cdots\right]$$

将上式右端展开，经整理可得

$$\ln M_{T_n}(t)=\frac{t^2}{2}+\cdots$$

省略项当 $n\to\infty$ 时趋于 0，所以

$$\lim_{n\to\infty}\ln M_{T_n}(t)=\frac{t^2}{2}$$

因此 $$\lim_{n\to\infty}M_{T_n}(t)=e^{\frac{t^2}{2}}$$

这正是 $N(0,1)$ 的矩母函数，由定理 4.2.5 就得证了 $F_n\xrightarrow{W}\Phi$。

特别地，上述定理中的 ξ_k 如果独立且均服从 0 — 1 分布，则 $\zeta_n=\xi_1+\xi_2+\cdots+\xi_n$ 亦是渐近正态分布。可是我们知道这时 ζ_n 是服从二项分布的，所以我们就有下面定理。

定理 4.3.2 （积分极限定理）设 ζ_n 服从参数为 (n,p) 的二项分布，$(0<p<1)$，则当 $n\to\infty$ 时，对任意区间 $[a,b]$，有

$$P\left(a\leqslant\frac{\zeta_n-np}{\sqrt{npq}}\leqslant b\right)\to\frac{1}{\sqrt{2\pi}}\int_a^b e^{-\frac{x^2}{2}}dx$$

积分极限定理给出了二项分布随机变量取值于某个区间的概率

的近似计算公式，即当 n 充分大时，

$$P\{a\leqslant\zeta_n\leqslant b\}=P\left\{\frac{a-np}{\sqrt{npq}}\leqslant\frac{\zeta_n-np}{\sqrt{npq}}\leqslant\frac{b-np}{\sqrt{npq}}\right\}$$

$$\doteq\int_{\frac{a-np}{\sqrt{npq}}}^{\frac{b-np}{\sqrt{npq}}}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}dx=\Phi\left(\frac{b-np}{\sqrt{npq}}\right)-\Phi\left(\frac{a-np}{\sqrt{npq}}\right)$$

$a<b$ 为任意的。

例4.3.1　已知大豆的虫食率为0.10，求在1000粒大豆中，虫食豆在100至200中间的概率。

解　1000粒大豆中的虫食豆的数目 ζ 服从参数为(1000，0.10)的二项分布，所以，

$$P\{100\leqslant\zeta\leqslant200\}\doteq\Phi\left(\frac{200-np}{\sqrt{npq}}\right)-\Phi\left(\frac{100-np}{\sqrt{npq}}\right)$$

$$=\Phi\left(\frac{200-1000\times0.10}{\sqrt{1000\times0.10\times0.90}}\right)-\Phi\left(\frac{100-1000\times0.10}{\sqrt{1000\times0.10\times0.90}}\right)$$

$$\doteq\Phi(10.5)-\Phi(0)\doteq0.5$$

例4.3.2　设某车间有200台车床，每台车床由于种种原因常常要停车。假定各台车床的停车或开车是相互独立的。若每台车床的开工率为0.6，开工时需要消耗的电能为 E，问发电厂至少要供给这个车间多少电能才能以99.9%的概率保证这个车间不致因为供电不足而影响生产。

解　用 ζ 表示200台车床中开车的台数。要解决这个问题，只要求出满足下面不等式的 k 就行了：

$$P\{\zeta\leqslant k\}\geqslant0.999,$$

因为上式表明，车间里有99.9%的可能开车台数不超过 k，因此，发电厂供给 k 台车床的电能就可以了。由积分极限定理可知

$$P\{\zeta\leqslant k\}\doteq\Phi\left(\frac{k-200\times0.6}{\sqrt{200\times0.6\times0.4}}\right)-\Phi\left(\frac{0-200\times0.6}{\sqrt{200\times0.6\times0.4}}\right)$$

解不等式

$$\Phi\left(\frac{k-120}{6.93}\right)-\Phi\left(\frac{-120}{6.93}\right)\geqslant 0.999$$

因为

$$\Phi\left(-\frac{120}{6.93}\right)\doteq\Phi(-17.3)\doteq 0$$

而

$$\Phi(3.1)=0.99903$$

所以

$$\frac{k-120}{6.93}\geqslant 3.1$$

$$k\geqslant 141.4\doteq 142$$

这就是说，发电厂只要供给这个车间 142E 的电能就足够了。此时，因为供电不足而停工影响生产的可能性小于 0.001 .

积分极限定理给出了二项分布随机变量当p，q都不很小时取值于某区间的概率的近似计算，普阿松定理给出了二项分布随机变量当 p 较小而 $\lambda=np$ 大小适中时的取值局部点的概率的近似计算. 那么，当 p，q 都不很小时，二项分布随机变量取值局部点的概率如何进行近似计算呢？下面的局部极限定理回答了这个问题.

定理4.3.3 （局部极限定理）设 ζ_n 服从参数为(n,p)的二项分布，$0<p<1$，以 x_k 表示$\frac{\zeta_n-E\zeta_n}{\sqrt{D\zeta_n}}=\frac{\zeta_n-np}{\sqrt{npq}}$的可能值$\frac{k-np}{\sqrt{npq}}$，$k=0,1,2\cdots,n$， $q=1-p$，则在 $n\to\infty$ 时，对一切使 x_k 在有限区间内的 k 值有

$$P\{\zeta_n=k\}\doteq\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{npq}}e^{-\frac{1}{2}x_k^2} \tag{4.3.1}$$

证明 因为 $x_k=\frac{k-np}{\sqrt{npq}}$，而对某一区间$[a,b]$有 $a\leqslant x_k\leqslant b$，则当$n\to\infty$时，

$$k=np+x_k\sqrt{npq}\to\infty$$

$$j=n-k=nq-x_k\sqrt{npq}\to\infty$$

引用数学分析中的斯提尔林公式

$$n! \doteq \sqrt{2\pi n}\, n^n e^{-n}$$

则有

$$P\{\zeta_n = k\} = \binom{n}{k} p^k q^{n-k} = \frac{n!}{k!\,j!} p^k q^{n-k}$$

$$\doteq \frac{\sqrt{2\pi n}\, n^n e^{-n}}{\sqrt{2\pi k}\, k^k e^{-k} \sqrt{2\pi j}\, j^j e^{-j}} p^k q^{n-k}$$

$$= \frac{1}{\sqrt{2\pi}} \sqrt{\frac{n}{kj}} \left(\frac{np}{k}\right)^k \left(\frac{nq}{j}\right)^j \tag{4.3.2}$$

但

$$\frac{kj}{n} = n \cdot \frac{k}{n} \cdot \frac{j}{n} = n\left(p + x_k\sqrt{\frac{pq}{n}}\right)\left(q - x_k\sqrt{\frac{pq}{n}}\right)$$

$$\doteq npq \tag{4.3.3}$$

为了计算后两项，先取其对数，

$$\ln\left(\frac{np}{k}\right)^k\left(\frac{nq}{j}\right)^j = -k\ln\left(\frac{k}{np}\right) - j\ln\left(\frac{j}{nq}\right)$$

$$= -(np + x_k\sqrt{npq})\ln\left(1 + x_k\sqrt{\frac{q}{np}}\right)$$

$$-(nq - x_k\sqrt{npq})\ln\left(1 - x_k\sqrt{\frac{p}{nq}}\right)$$

将对数函数按马克洛林级数展开，就有

$$\ln\left(\frac{np}{k}\right)^k\left(\frac{nq}{j}\right)^j = -(np + x_k\sqrt{npq})\left[x_k\sqrt{\frac{q}{np}}\right.$$

$$\left.-\frac{1}{2}\frac{q}{np}x_k^2 + \cdots\right] - (nq - x_k\sqrt{npq})$$

$$\left[-x_k\sqrt{\frac{p}{nq}} - \frac{1}{2}\frac{p}{nq}x_k^2 + \cdots\right]$$

$$= -x_k\sqrt{npq} - qx_k^2 + \frac{1}{2}qx_k^2 + \cdots$$

$$+ x_k\sqrt{npq} - px_k^2 + \frac{1}{2}px_k^2 + \cdots$$

$$= -\frac{1}{2}x_k^2 + \cdots$$

从而
$$\left(\frac{np}{k}\right)^k\left(\frac{nq}{j}\right)^j \doteq e^{-\frac{1}{2}x_k^2} \qquad (4.3.4)$$

把(4.3.3)、(4.3.4)代入(4.3.2)即得

$$P\{\xi_n = k\} \doteq \frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{npq}}e^{-\frac{1}{2}x_k^2}$$

例4.3.3 在上一章曾对参数为(40,0.25)的二项分布用普阿松定理进行了近似计算，结果表明误差较大。现在用局部极限定理给出同一问题的近似计算，结果

k	按二项分布	按局部极限定理	按普阿松定理
5	0.0272	0.0273	0.0378
6	0.0529	0.0502	0.0631
7	0.0857	0.0795	0.0901
8	0.1179	0.1115	0.1126
9	0.1398	0.1360	0.1251
10	0.1444	0.1456	0.1251
11	0.1312	0.1360	0.1137
12	0.1057	0.1115	0.0948
13	0.0759	0.0795	0.0729
14	0.0488	0.0502	0.0521
15	0.0282	0.0273	0.0347

可以看出，对p不太小的情况，局部极限定理给出的近似程度确实比用普阿松定理好些。当n再大些时，这个区别会更加明显。

例4.3.4 设次品率为0.005，求在10000件产品中，（1）有40的件次品的概率；（2）次品不多于70件的概率。

解 10000件产品中的次品数ξ服从参数为(10000，0.005)的二项分布，（1）用局部极限定理

$$P\{\xi=40\}\doteq\frac{1}{\sqrt{10000\times0.995\times0.005}\sqrt{2\pi}}$$

$$e^{-\frac{1}{2}\left(\frac{40-10000\times0.005}{\sqrt{10000\times0.995\times0.005}}\right)^2}$$

$$\doteq\frac{1}{7.05}\frac{1}{\sqrt{2\pi}}e^{-\frac{(1.42)^2}{2}}=\frac{1}{7.05}\varphi(1.42)$$

查表得 $P\{\xi=40\}\doteq\frac{0.1456}{7.05}\doteq0.0206$

（2）

$$P\{0\leqslant\xi\leqslant70\}=\Phi\left(\frac{70-10000\times0.005}{\sqrt{10000\times0.995\times0.005}}\right)$$

$$-\Phi\left(\frac{0-10000\times0.005}{\sqrt{10000\times0.995\times0.005}}\right)$$

$$=\Phi(2.84)-\Phi(-7.09)$$

$$=\Phi(2.84)+\Phi(7.09)-1\doteq0.9977$$

上面我们所讨论的中心极限定理都是同分布中心极限定理，那么对于不是具有相同分布的独立随机变量序列是否也有类似的极限定理呢？下面所要介绍的李雅普诺夫定理就是一种这样的中心极限定理。为了给出定理的证明，我们先引出一个关于随机变量绝对矩之间关系的不等式——李雅普诺夫不等式。

引理（李雅普诺夫不等式）若随机变量 ξ 存在 n 阶绝对矩，则对于 $k=1,2,\cdots,n-1$，有

$$a_k^{1/k}\leqslant a_{k+1}^{1/k+1}$$

其中 a_k 为 ξ 的 k 阶绝对矩。

证明 我们仅就 ξ 为连续型随机变量场合给出证明，至于离散情形的证明与连续情形类似。设 $f(x)$ 为 ξ 的分布密度，对于任意的两个实数 u，v，考虑下列积分

$$0\leqslant\int_{-\infty}^{\infty}\left[u|x|^{\frac{k-1}{2}}+v|x|^{\frac{k+1}{2}}\right]^2f(x)dx$$

$$= \int_{-\infty}^{\infty}\left[u^2|x|^{k-1} + 2uv|x|^{k} + v^2|x|^{k+1}\right]f(x)dx$$

$$= \alpha_{k-1}u^2 + 2\alpha_k uv + \alpha_{k+1}v^2$$

由于此二次式永不变号，故

$$\alpha_k^2 - \alpha_{k-1}\alpha_{k+1} \leqslant 0$$

即 $$\alpha_k^2 \leqslant \alpha_{k-1}\alpha_{k+1}$$

把得到的不等式两端各自自乘 k 次，便得到

$$\alpha_k^{2k} \leqslant \alpha_{k-1}^k \alpha_{k+1}^k$$

依次令 $k=1,2,\cdots,n-1$，可得

$$\alpha_1^2 \leqslant \alpha_0\alpha_2 \qquad \left(\alpha_0 = \int_{-\infty}^{\infty}|x|^0 f(x)dx = 1\right)$$

$$\alpha_2^4 \leqslant \alpha_1^2\alpha_3^2$$

$$\alpha_3^6 \leqslant \alpha_2^3\alpha_4^3$$

$$\cdots\cdots\cdots\cdots$$

$$\alpha_{n-1}^{2(n-1)} \leqslant \alpha_{n-2}^{n-1}\alpha_n^{n-1}$$

将这些不等式中前 k 个的左端及右端分别相乘，并注意到 $\alpha_0=1$，便得到对于 $k=1,2,\cdots,n-1$ 均成立的下列不等式

$$\alpha_k^{k+1} \leqslant \alpha_{k+1}^k$$

再将此不等式开 $k(k+1)$ 次方，即得到引理欲证结论。

*定理4.3.4（李雅普诺夫）设$\{\xi_n\}$是相互独立随机变量序列，又设 a_k，$b_k \neq 0$，c_k，d_k 分别为随机变量 ξ_k 的均值，标准差，三阶中心矩和三阶绝对中心矩。令 $B_n = \sqrt{\sum_{k=1}^n b_k^2}$，$D_n = \sqrt[3]{\sum_{k=1}^n d_k}$。

如果它们满足条件 $$\lim_{n\to\infty}\frac{D_n}{B_n} = 0 \tag{4.3.5}$$

那么随机变量 $\zeta_n = \sum_{k=1}^n \frac{\xi_k - a_k}{B_n}$

的分布函数 $F_n(x)$就收敛于 $\Phi(x)$，其中 $\Phi(x)$为 $N(0,1)$随机变量的分布函数。

---

* 定理4.3.4不是必读内容以下凡有*的部分皆为非必读内容。

证明　令　$\eta_k=\frac{\xi_k-a_k}{B_n}$

以 $M_k(t)$表示 ξ_k-a_k的矩母函数，因$E(\xi_k-a_k)=0$，且 b_k^2 和 c_k均存在，所以可把 $M_k(t)$展成马克劳林级数

$$M_k(t)=1+\frac{b_k^2t^2}{2}+\frac{c_kt^3}{6}+o(c_kt^3)$$

于是做为 ξ_k-a_k 的线性函数 η_k 的矩母函数

$$M_{\eta_k}(t)=M_k\left(\frac{t}{B_n}\right)=1+\frac{b_k{}^2t^2}{2B_n^2}+\frac{c_kt^3}{6B_n^3}+o\left(\frac{c_kt^3}{B_n^3}\right)$$

$$=1+u_k$$

由李雅普诺夫不等式，有 $b_k\leqslant\sqrt[3]{d_k}$，于是从(4.3.5)式可推得，对任何有限 t，有

$$\lim_{n\to\infty}\left|\frac{b_k^2t^2}{2B_n^2}\right|\leqslant\lim_{n\to\infty}\frac{\sqrt[3]{d_k}}{2B_n^2}t^2\leqslant\lim_{n\to\infty}\frac{1}{2}\frac{D_n^2}{B_n}t^2=0$$

$$\lim_{n\to\infty}\left|\frac{c_kt^3}{6B_n^3}\right|\leqslant\lim_{n\to\infty}\frac{d_k}{6B_n^3}|t^3|\leqslant\lim_{n\to\infty}\frac{1}{6}\frac{D_n^3}{B_n^3}|t^3|=0.$$

而$o\left(\frac{c_kt^3}{B_n^3}\right)$是关于$\frac{c_kt^3}{B_n^3}$的更高级无穷小量，故有$\lim\limits_{n\to\infty}o\left(\frac{c_kt^3}{B_n^3}\right)=0$。

综上所述可得$\lim\limits_{n\to\infty}u_k=0$。这里的收敛关于$k$是均匀的。因此对任意 t，总存在一个数 $N=N(t)$，使得对于 $n>N$，及所有 $k\leqslant n$，有$|u_k|<\frac{1}{2}$。考虑 $M_{\eta_k}(t)$ 的对数的马克劳林展开

$$\ln M_{\eta_k}(t)=\ln(1+u_k)$$

$$=u_k-\frac{u_k^2}{2}+\frac{u_k^3}{3}-\cdots$$

$$=u_k-\frac{u_k^2}{2}\left(1-\frac{2}{3}u_k+\frac{2}{4}u_k^2-\cdots\right)$$

$$=u_k-\frac{u_k^2}{2}v_k$$

这里 v_k 满足不等式

$$|v_k| \leqslant 1 + \frac{2}{3}|u_k| + \frac{2}{4}|u_k|^2 + \cdots < 1 + |u_k| + |u_k|^2 + \cdots$$

$$< 1 + \frac{1}{2} + \frac{1}{4} + \cdots = 2$$

所以我们有

$$\ln M_{\eta_k}(t) = u_k + \theta_k u_k^2$$

其中 $\theta_k = -\frac{v_k}{2}$，而 $|\theta_k| < 1$。

设 $M_{\zeta_n}(t)$ 为 ζ_n 的矩母函数，则

$$M_{\zeta_n}(t) = \prod_{k=1}^{n} M_{\eta_k}(t)$$

因此

$$\ln M_{\zeta_n}(t) = \sum_{k=1}^{n} \ln M_{\eta_k}(t)$$

$$= \sum_{k=1}^{n} (u_k + \theta_k u_k^2) \tag{4.3.6}$$

又
$$\sum_{k=1}^{n} u_k = \frac{t^2}{2} + \sum_{k=1}^{n} \frac{c_k t^3}{6B_n^3} + \sum_{k=1}^{n} o\left(\frac{c_k t^3}{B_n^3}\right)$$

因为对于任何 t，

$$\lim_{n\to\infty} \left| \sum_{k=1}^{n} \frac{c_k t^3}{6B_n^3} \right| \leqslant \lim_{n\to\infty} \sum_{k=1}^{n} \frac{d_k |t^3|}{6B_n^3}$$

$$\leqslant \lim_{n\to\infty} \frac{D_n^3 |t^3|}{6B_n^3} = 0$$

$$\lim_{n\to\infty} \sum_{k=1}^{n} o\left(\frac{c_k t^3}{B_n^3}\right) = \lim_{n\to\infty} \sum_{k=1}^{n} \left\{ \frac{c_k t^3}{6B_n^3} \left[o\left(\frac{c_k t^3}{B_n^3}\right) \Big/ \frac{c_k t^3}{6B_n^3} \right] \right\}$$

$$= 0$$

所以有 $\lim\limits_{n\to\infty} \sum\limits_{k=1}^{n} u_k = \frac{t^2}{2}$

下面来估计和

$$\sum_{k=1}^{n} u_k^2 = \sum_{k=1}^{n} \left[\frac{b_k^2 t^2}{2B_n^2} + \frac{c_k t^3}{6B_n^3} + o\left(\frac{c_k t^3}{B_n^3}\right) \right]^2$$

的极限值。由李雅普诺夫不等式，有

$$\lim_{n\to\infty}\sum_{k=1}^{n}\frac{b_k^4t^4}{4B_n^4}\leqslant\lim_{n\to\infty}\sum_{k=1}^{n}\frac{\sqrt[3]{d_k^4}}{4B_n^4}t^4=\lim_{n\to\infty}\sum_{k=1}^{n}\frac{d_k\sqrt[3]{d_k}}{4B_n^3B_n}t^4$$

$$\leqslant\lim_{n\to\infty}\sum_{k=1}^{n}\frac{d_k}{4B_n^3}t^4=\lim_{n\to\infty}\frac{D_n^3}{4B_n^3}t^4=0$$

$$\lim_{n\to\infty}\left|\sum_{k=1}^{n}\left(\frac{c_kt^3}{6B_n^3}\right)^2\right|\leqslant\lim_{n\to\infty}\sum_{k=1}^{n}\left|\frac{c_kt^3}{6B_n^3}\right|^2\leqslant\lim_{n\to\infty}\sum_{k=1}^{n}\frac{d_k^2t^6}{36B_n^6}$$

$$=\lim_{n\to\infty}\frac{D_n^6t^6}{36B_n^6}=0$$

同样，有 $\lim\limits_{n\to\infty}\left|\sum\limits_{k=1}^{n}\frac{b_k^2c_kt^3}{6B_n^5}\right|=0$

$$\lim_{n\to\infty}\left|\sum_{k=1}^{n}\frac{b_k^2t^2}{B_n^2}o\left(\frac{c_kt^3}{B_n^3}\right)\right|=0$$

$$\lim_{n\to\infty}\left|\sum_{k=1}^{n}\frac{c_kt^3}{3B_n^3}o\left(\frac{c_kt^3}{B_n^3}\right)\right|=0$$

综上所述，并注意到 $k\leqslant n$， $|\theta_k|<1$我们有

$$\lim_{n\to\infty}\sum_{k=1}^{n}\theta_ku_k^2=0$$

由（4.3.6)式，则有$\lim\limits_{n\to\infty}\ln M_{\zeta_n}(t)=\frac{t^2}{2}$

即 $\lim\limits_{n\to\infty}M_{\zeta_n}(t)=e^{\frac{t^2}{2}}$

由矩母函数的连续性唯一性定理则可推得 $\zeta_n\xrightarrow{L}\zeta$（其中 $\zeta\sim N(0,1)$)。即ζ_n 的分布函数 $F_n(x)\xrightarrow{W}\Phi(x)$。

现在转回来探讨随机变量和渐近正态分布的一般化的条件，或者说给出条件更一般化的中心极限定理。

从直观上或从经验上我们已经知道，大量的、相互独立的、每

一个对总体的影响都"均匀地小"的随机变量的总效果是呈正态分布的。下面我们就从理论上来论述它。

设 $\xi_1, \xi_2, \cdots, \xi_n, \cdots$ 是一个相互独立的随机变量序列，它们具有有限的均值和方差：

$$a_k = E\xi_k, \quad b_k^2 = D\xi_k \quad (k=1,2,\cdots)$$

记 $$B_n^2 = \sum_{k=1}^{n} b_k^2$$

B_n^2 是 $\xi_1+\xi_2+\cdots+\xi_n$ 的方差，其标准差为 B_n。每一个 ξ_k 的离差是 $\xi_k - a_k$，它对总体的相对影响或"贡献"可用 $\frac{\xi_k - a_k}{B_n}$ 表示，所以总影响或总贡献为和数

$$\zeta_n = \sum_{k=1}^{n} \frac{\xi_k - a_k}{B_n} \tag{4.3.7}$$

那么什么叫每一个对总体影响都"均匀地小"呢？考虑每一个影响的绝对值 $\frac{|\xi_k - a_k|}{B_n}$，所谓每个都"均匀地小"，粗略地说，就是所有这些影响最大者都不能超过任何一个指定的正数 τ，即要求：

$$P\left(\max_{1\leqslant k\leqslant n} \frac{|\xi_k - a_k|}{B_n} > \tau\right) = P\left(\bigcup_{k=1}^{n} (|\xi_k - a_k| > \tau B_n)\right)$$

$$= P\left(\bigcup_{k=1}^{n} A_k\right) \xrightarrow[n\to\infty]{} 0$$

其中 $A_k = (|\xi_k - a_k| > \tau B_n) \quad (k=1,2,\cdots,n)$。另一方面，

$$P\left(\bigcup_{k=1}^{n} A_k\right) \leqslant \sum_{k=1}^{n} P(A_k) = \sum_{k=1}^{n} \int_{|x-a_k|>\tau B_n} dF_k(x)^*$$

$$\leqslant \frac{1}{(\tau B_n)^2} \sum_{k=1}^{n} \int_{|x-a_k|>\tau B_n} (x-a_k)^2 dF_k(x)$$

---

* 积分 $\int_{|x-a_k|>\tau B_n} dF_k(x)$ 为斯梯阶积分，当随机变量是连续型或离散型时，此积分分别变成密度函数的积分或分布列求和。

$$=\frac{1}{\tau^2 B_n^2}\sum_{k=1}^{n}\int_{|x-a_k|>\tau B_n}(x-a_k)^2dF_k(x)$$

所以只要对任何 $\tau>0$，有

$$\lim_{n\to\infty}\frac{1}{B_n^2}\sum_{k=1}^{n}\int_{|x-a_k|>\tau B_n}(x-a_k)^2dF_k(x)=0 \qquad (4.3.8)$$

就可以保证总和(4.3.7)式中加项"均匀地小"．上述条件(4.3.8)称为林德贝格条件．林德贝格证明了条件(4.3.8)是和数(4.3.7)的分布函数趋于正态分布的充分条件，下面我们只叙述此结论而不加证明。

定理4.3.5 (*Linderberg*) 设随机变量序列$\{\xi_k\}$相互独立满足林德贝格条件(4.3.8)，则有

$$\lim_{n\to\infty}P\left(\frac{1}{B_n}\sum_{k=1}^{n}(\xi_k-a_k)\leqslant x\right)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x}e^{-\frac{z^2}{2}}dz$$

可以验证定理(4.3.1)和定理(4.3.4)均为定理 (4.3.5) 的特殊情形．显然只要分别验证林德贝格条件都满足就行了．事实上，在定理 (4.3.1) 的条件下， $B_n=\sqrt{n}\,\sigma$，而

$$\frac{1}{B_n^2}\sum_{k=1}^{n}\int_{|x-a_k|>\tau B_n}(x-a_k)^2dF_k(x)$$

$$=\frac{1}{n\sigma^2}\ n\int_{|x-a_k|>\tau\sigma\sqrt{n}}(x-a_k)^2dF_k(x)$$

由于 $0<\sigma^2<\infty$，所以上式右边的积分当 $n\to\infty$ 时趋于零，即林德贝格条件得到满足．在定理 (4.3.4) 条件下，

$$\frac{1}{B_n^2}\sum_{k=1}^{n}\int_{|x-a_k|>\tau B_n}(x-a_k)^2dF_k(x)$$

$$\leqslant\frac{1}{B_n^2(\tau B_n)}\sum_{k=1}^{n}\int_{|x-a_k|>\tau B_n}|x-a_k|^3dF_k(x)$$

$$\leqslant\frac{1}{\tau B_n^3}\sum_{k=1}^{n}\int_{-\infty}^{\infty}|x-a_k|^3dF_k(x)$$

$$= \frac{1}{\tau} \cdot \frac{D_n^3}{B_n^3} \to 0 \quad (\text{当} n \to \infty \text{时})$$

§4.4 大 数 定 理

在概率论中讨论一个命题的真伪时，一般都要指出该命题成立的可能性的大小，这里所谓一个恒真的命题往往就是指该命题成立的可能性（概率）为 1，当然不真命题就是指其成立的概率为 0．自然，在讨论一个命题时，最理想的结论就是指出在什么条件下，该命题成立的概率为1（或为 0）．比如在天气预报中，如果能得出什么时候下雨的概率为 1 的结论，就是最为理想的结论．也就说实际中需要我们讨论渐近于 0 — 1 律的极限定理，这就是我们本节要介绍的所谓大数定律．

（一）切贝晓夫不等式

为了讨论大数定律，我们先证明一个重要的不等式．

切贝晓夫不等式　对任何具有有限方差的随机变量 ξ，都有

$$P(|\xi - E\xi| \geqslant \varepsilon) \leqslant \frac{D\xi}{\varepsilon^2}.$$

其中 ε 为任一正数．

证明　我们对 ξ 是离散型随机变量和连续型随机变量两种不同情形分开来证明．若 ξ 是离散型随机变量，假设它的分布列是

$$\begin{pmatrix} x_1, & x_2, & \cdots, & x_n, & \cdots \\ p_1, & p_2, & \cdots, & p_n, & \cdots \end{pmatrix}$$

根据方差的定义

$$D\xi = \sum_{n=1}^{\infty} (x_n - E\xi)^2 p_n$$

将上式中的和分成两个部分，

$$D\xi = \sum_{|x_n - E\xi| \geqslant \varepsilon} (x_n - E\xi)^2 p_n + \sum_{|x_n - E\xi| < \varepsilon} (x_n - E\xi)^2 p_n$$

其中第一部分是对满足 $|x_n - E\xi| \geqslant \varepsilon$ 的一切项求和，第二部分是对其余的项求和．因为和中每项都是非负的，去掉第二部分，这个和就只能减小，所以，

$$D\xi \geqslant \sum_{|x_n - E\xi| \geqslant \varepsilon} (x_n - E\xi)^2 p_n$$

如果在其中的每一项我们再用较小的量 ε^2 代替 $(x_n - E\xi)^2$，和就更加减少，因此，

$$D\xi \geqslant \varepsilon^2 \sum_{|x_n - E\xi| \geqslant \varepsilon} p_n$$

位于上式右端的和，是随机变量 ξ 取满足 $|x_n - E\xi| \geqslant \varepsilon$ 的值 x_n 的概率的和，也就是概率 $P(|\xi - E\xi| \geqslant \varepsilon)$，所以，我们得到

$$D\xi \geqslant \varepsilon^2 P(|\xi - E\xi| \geqslant \varepsilon),$$

即

$$P(|\xi - E\xi| \geqslant \varepsilon) \leqslant \frac{D\xi}{\varepsilon^2}.$$

当 ξ 是连续型随机变量时，假设它的分布密度是 $p(x)$，根据和离散型随机变量时相类似的分析，可得

$$\begin{aligned}
D\xi &= \int_{-\infty}^{\infty} (x - E\xi)^2 p(x) dx \\
&= \int_{|x - E\xi| \geqslant \varepsilon} (x - E\xi)^2 p(x) dx \\
&\qquad + \int_{|x - E\xi| < \varepsilon} (x - E\xi)^2 p(x) dx \\
&\geqslant \int_{|x - E\xi| \geqslant \varepsilon} (x - E\xi)^2 p(x) dx \\
&\geqslant \varepsilon^2 \int_{|x - E\xi| \geqslant \varepsilon} p(x) dx = \varepsilon^2 P(|\xi - E\xi| \geqslant \varepsilon),
\end{aligned}$$

即

$$P(|\xi - E\xi| \geqslant \varepsilon) \leqslant \frac{D\xi}{\varepsilon^2}. \qquad (4.4.1)$$

我们知道，方差 $D\xi$ 描述了随机变量 ξ 对它的均值的离散程度，更确切地说，是描述 ξ 对 $E\xi$ 的平均离散程度．有时，我们还需要知道 ξ 对 $E\xi$ 产生大偏差的可能性多大，即需要计算 $|\xi - E\xi|$ 不小

于某一事先给定数ε的概率 $P(|\xi-E\xi|\geqslant\varepsilon)$ 。切贝晓夫不等式说明了方差愈小 $P(|\xi-E\xi|\geqslant\varepsilon)$愈小，因而，它进一步说明了方差确实能描述随机变量对它均值的离散程度。

(4.4.1)还可改写成

$$P(|\xi-E\xi|\geqslant\varepsilon\sqrt{D\xi})\leqslant\frac{1}{\varepsilon^2} \tag{4.4.2}$$

(4.4.2) 式给出了随机变量偏离均值为方差的某一倍数的可能性大小的估计。比如，若 $\varepsilon=\frac{3}{2}$，由 (4.4.2) 式我们有

$$P\left\{|\xi-E\xi|\geqslant\frac{3}{2}\sqrt{D\xi}\right\}\leqslant\frac{4}{9}\doteq0.44$$

（二）大数定律

借助于切贝晓夫不等式，我们可以导出一个重要而常用的大数定律。

定理4.4.1（切贝晓夫定理） 如果ξ_1，ξ_2，ξ_n，… 是两两相互独立的随机变量序列，每一随机变量都有有限的方差，并且不大于同一常数 c

$$D\xi_1\leqslant c,\ D\xi_2\leqslant c,\ \cdots,\ D\xi_n\leqslant c,\ \cdots$$

则对任意的 $\varepsilon>0$，都有

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i-\frac{1}{n}\sum_{i=1}^{n}E\xi_i\right|<\varepsilon\right\}=1$$

$\left(\text{或等价说，}\frac{1}{n}\sum_{i=1}^{n}\xi_i\xrightarrow{P}\frac{1}{n}\sum_{i=1}^{n}E\xi_i\right)$.

证明 考虑随机变量$\frac{1}{n}\sum_{i=1}^{n}\xi_i$，它的均值

$$E\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right)=\frac{1}{n}\sum_{i=1}^{n}E\xi_i,$$

又因为$\{\xi_i\}$是两两独立的，方差

$$D\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}D\xi_i\leqslant\frac{c}{n}$$

由切贝晓夫不等式得

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i-\frac{1}{n}\sum_{i=1}^{n}E\xi_i\right|\geqslant\varepsilon\right\}\leqslant\frac{D\left(\frac{1}{n}\sum_{i=1}^{n}\xi_i\right)}{\varepsilon^2}\leqslant\frac{c}{n\varepsilon^2}$$

从而

$$P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i-\frac{1}{n}\sum_{i=1}^{n}E\xi_i\right|<\varepsilon\right\}=1-P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i-\right.\right.$$

$$\left.\left.-\frac{1}{n}\sum_{i=1}^{n}E\xi_i\right|\geqslant\varepsilon\right\}\geqslant1-\frac{c}{n\varepsilon^2}$$

所以

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i-\frac{1}{n}\sum_{i=1}^{n}E\xi_i\right|<\varepsilon\right\}\geqslant1$$

由于概率不能大于 1，所以有

$$\lim_{n\to\infty}P\left\{\left|\frac{1}{n}\sum_{i=1}^{n}\xi_i-\frac{1}{n}\sum_{i=1}^{n}E\xi_i\right|<\varepsilon\right\}=1$$

切贝晓夫定理告诉我们，尽管随机变量序列

$\xi_1,\ \xi_2,\ \cdots,\ \xi_n,\cdots$

中的每一个，由于种种偶然因素影响，在各次试验中所取得的值是以偶然的方式变化着，然而，在某些条件下，只要 n 足够大，n 个随机变量的算术平均就服从一个完全确定的规律，即这个算术平均只能围绕一个固定常数取值，它和这常数有显著偏差的可能性是很小的。这就再一次说明，必然性和偶然性是对立的统一。

由切贝晓夫定理可以推出

定理4.4.2（贝努利定理） 设 ζ_n 是 n 次重复独立试验中事件 A 出现的次数，而 p 是事件 A 在每次试验中出现的概率，则对任意的 $\varepsilon>0$，都有

$$\lim_{n\to\infty}P\left\{\left|\frac{\zeta_n}{n}-p\right|<\varepsilon\right\}=1$$

证明 以 μ_k 表示在第 k 次试验中事件 A 出现的次数，显然

$$\xi_n = \mu_1 + \mu_2 + \cdots + \mu_n$$

每个 μ_k 都是具有同一分布

$$\begin{pmatrix} 0, & 1 \\ q, & p \end{pmatrix} \qquad q = 1 - p$$

的随机变量，

$$E\mu_k = p \qquad D\mu_k = pq$$

对随机变量序列$\{\mu_k\}$应用切贝晓夫定理即得贝努利定理。

贝努利定理确切地揭示了频率与概率之间的关系，就是当试验次数很大时，事件出现的频率与概率有较大偏差的可能性很小。这就是我们经验上所熟悉的"频率稳定性"的理论上的证明，是我们用频率估计概率的依据。

最后，由切贝晓夫定理还可以推出下面的定理

定理4.4.3（平均数法则）如果 $\xi_1, \xi_2, \cdots, \xi_n, \cdots$ 是两两相互独立的随机变量序列，每个随机变量都服从相同分布，且具有有限方差，

$$E\xi_n = a \qquad D\xi_n < \infty \qquad n = 1, 2, \cdots$$

则对任意的 $\varepsilon > 0$，有

$$\lim_{n \to \infty} P\left\{ \left| \frac{1}{n} \sum_{i=1}^{n} \xi_i - a \right| < \varepsilon \right\} = 1$$

在实际问题中经常应用平均数法则，例如当测定某一物理量时，我们总是愿意多测几次，然后取其算术平均作为这个物理量的近似值，并且经验告诉我们，一般说来，测的次数越多，近似的程度越好，这正是平均数法则在起作用。

我们看到，在贝努利定理和平均数法则中相应的随机变量序列都是独立同分布的。一般来说，如果已知随机变量序列是独立同分布的，实际上用不着涉及方差的性质，就可以断言它是否服从大数定律。

定理4.4.4（辛钦定理） 设 $\{\xi_n\}$ 独立同分布且具有有限均值

$\mu=E\xi_1$，则当 $n\to\infty$时有

$$\frac{1}{n}\sum_{i=1}^{n}\xi_i\xrightarrow{P}\mu$$

证明　我们仅就ξ_k的矩母函数存在的情形给出定理的证明.由于$\xi_1,\xi_2,\cdots,\xi_n$同分布，故有相同的矩母函数，设为$M(t)$。又因均值存在，所以$M(t)$可展成

$$\begin{aligned}M(t)&=M(0)+\frac{M'(0)}{1!}t+o(t)\\&=1+\mu t+o(t)\end{aligned}$$

$\frac{1}{n}\sum_{i=1}^{n}\xi_i$的矩母函数为

$$\left[M\left(\frac{t}{n}\right)\right]^n=\left[1+\mu\frac{t}{n}+o\left(\frac{t}{n}\right)\right]^n$$

对固定的t，$\left[M\left(\frac{t}{n}\right)\right]^n\longrightarrow e^{\mu t}\qquad(n\to\infty)$

$e^{\mu t}$ 正是退化分布$\begin{pmatrix}\mu\\1\end{pmatrix}$的矩母函数．由矩母函数连续性定理及定理4.1.6则可推出$\frac{1}{n}\sum_{i=1}^{n}\xi_i\xrightarrow{P}\mu$.

习　题

1　设ξ的分布密度为

$f(x)=2x$，$0\leqslant x\leqslant1$。

(a)求ξ的矩母函数

(b)利用矩母函数计算$E\xi$和$D\xi$。

2　设ξ具有如下的分布密度

$f(x)=\lambda e^{-\lambda(x-a)}$，$x>a$

(它是熟知的两个参数的指数分布)。

(a)求ξ的矩母函数。

(b)利用矩母函数求 $E\xi$ 和 $D\xi$。

3　若 $\xi \sim N(\mu_\xi, \sigma_\xi^2)$

$\eta \sim N(\mu_\eta, \sigma_\eta^2)$

且 ξ 与 η 相互独立。试利用矩母函数证明 $\zeta = a\xi + b\eta$ 也是正态分布的，其中 a，b均为常数。

4　设随机变量 ξ 的矩母函数为

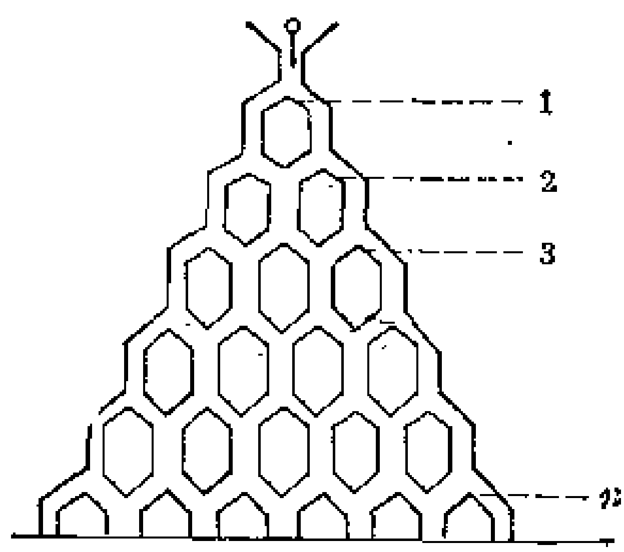
$M_\xi(t) = (0.4e^t + 0.6)^8$

(a)随机变量 $\eta = 3\xi + 2$ 的矩母函数是什么？

(b)求 $E\xi$。

5　设 ξ_1 和 ξ_2 相互独立，且分别服从参数为 λ_1 和 λ_2 的 *Poisson* 分布，试用矩母函数方法求 $\zeta = \xi_1 + \xi_2$ 的分布。

6　在图（4—6题）入口处放进一球，求它落入第 n 层各点的概率，当 $n \to \infty$ 时，其概率为多少。试分析一下当不断地往入口处放球时，球的分布形状。



图（4—6题）

7　从发芽率为0.95的一批种子里随机取 400 粒，试求其不发芽的种子不多于25粒的概率。

8　假设 $\xi_i, i = 1, 2, \cdots, 50$，独立同 *Poisson* 分布，其参数为 $\lambda = 0.03$。令 $\zeta = \xi_1 + \xi_2 + \cdots + \xi_{50}$，

(a)利用中心极限定理求

$P(\zeta \geqslant 3)$。

(b)把(a)的结果与这个概率的精确值进行比较。

9　从装有 3 个白球和 1 个黑球的箱中，有放回地取 n 个球，设 ξ 是白球出现的次数，问 n 需要多大才能使得

$$P\left(\left|\frac{\zeta}{n}-p\right|\leqslant 0.001\right)=0.9964$$

其中 p 是每次取到白球的概率。

10　设$\{\xi_i\}$独立且同在$[0,T]$区间上均匀分布，试求 $\zeta_2=\xi_1+\xi_2$，$\zeta_3=\xi_1+\xi_2+\xi_3$ 的分布密度、均值和方差，并写出与所求出的均值、方差相同的正态分布的密度函数，画出ζ_2，ζ_3及相应的正态分布的分布密度图加以比较。

11　设$\{\xi_i\}$独立同分布 $N(0,1)$，求随机变量$W_n=\sqrt{n}\dfrac{\sum\limits_{i=1}^{n}\xi_i}{\sum\limits_{i=1}^{n}\xi_i^2}$的极限分布。

12　设随机变量 ξ_n 的分布列为$P(\xi_n=1)=\dfrac{1}{n}$，$P(\xi_n=0)=1-\dfrac{1}{n}$，$\xi$ 是退化分布$\begin{pmatrix}0\\1\end{pmatrix}$随机变量，求证 $\xi_n\xrightarrow{L}\xi$。

13　设 $\xi_\lambda\sim p(\lambda)$，$\xi\sim N(0,1)$，求证$\dfrac{\xi_\lambda-\lambda}{\sqrt{\lambda}}\xrightarrow{L}\xi$。

14　设$\{\xi_i\}$是两两不相关随机变量序列，$E\xi_i=\mu_i$，$D\xi_i=\sigma_i^2$，$i=1,2,\cdots$。假如当 $n\to\infty$时，$\sum\limits_{i=1}^{n}\sigma_i^2\to\infty$，则当 $n\to\infty$时有

$$\sum_{i=1}^{n}\frac{\xi_i-\mu_i}{\sum\limits_{i=1}^{n}\sigma_i^2}\xrightarrow{P}0$$

学 习 指 导

(一) 本章概要

矩母函数既是讨论随机变量序列极限分布的一个有力工具同

时本身又有它的明确的概率涵义．所以必须首先掌握住它的定义，记住各重要随机变量的矩母函数以及弄清它与各阶矩之间的关系．其次还必须理解唯一性定理和连续性定理的意义和用途．

为了研究极限定理，还须搞清诸收敛性，它们的涵义、性质及它们之间的关系．

中心极限定理的理论意义很大，要掌握同分布中心极限定理的内容及其证明方法．中心极限定理给出二项分布近似计算的新途径，要求能够正确应用积分极限定理和局部极限定理对二项分布进行近似计算．

要理解大数定律的理论价值及切贝晓夫不等式和切贝晓夫定理的证明．

（二）内容说明

1 关于矩母函数的定义．从矩母函数的定义可以看出，它正好是密度函数$f(x)$的拉普拉斯变换．与拉氏变换作用相仿的还有富氏变换．如果对函数$f(x)$进行富氏变换，就得到一个新的函数

$$C(t)=\int_{-\infty}^{\infty}e^{itx}f(x)dx$$

我们称之为特征函数，特征函数是一个实变量复值函数，由于$|e^{itx}|=1$，所以它对一切实数t都是有意义的．对于特征函数来说，同样有矩母函数的唯一性和连续性定理，它也是研究随机变量渐近分布的有效工具．特征函数比起矩母函数来，优点是这个函数对一切t都有意义，缺点是特征函数要涉及到复数的工具且不像矩母函数那样具有直观的概率意义．

2 关于依分布收敛和依概率收敛．虽然它们讲的是随机变量的收敛问题，但实际上涉及的却分别是序列$\{F_n(x)\}$和$P(|\xi_n-\xi|>\varepsilon)$的收敛问题，所以在理解它们的涵义时，要注意和数学分析中通常碰到的收敛的概念区别开来．

3 关于同分布中心极限定理的证明．不管是应用矩母函数还

是应用特征函数作为工具来证明随机变量的极限分布时，同分布中心极限定理证明中所给出的证明步骤是十分典型的，可以把它归纳成如下几步：

（1）求出相应的随机变量序列的矩母函数，把随机变量的分布的讨论化成矩母函数的讨论.

（2）对矩母函数取对数.

（3）进行马克劳林展开，求极限.

（4）对极限求反对数，就得到极限的矩母函数.

例4.1 设 $\xi_n \sim \chi^2(n)$，$n=1,2,\cdots$. 求 ξ_n/n^2 的极限分布.

因 $\xi_n \sim \chi^2(n)$，所以 ξ_n 的矩母函数为 $M_n(t)=\left(\dfrac{1}{1-2t}\right)^{n/2}$.

令 $S_n=\xi_n/n^2$，则 S_n 的矩母函数为

$$M_{S_n}(t)=\left(\frac{1}{1-\dfrac{2t}{n^2}}\right)^{n/2}=(1-2t/n^2)^{-n/2}$$

取对数 $\quad \ln M_{S_n}(t)=-\dfrac{n}{2}\ln(1-2t/n^2)$

$$=-\frac{n}{2}\left(-\frac{2t}{n^2}-\frac{(-2t/n^2)^2}{2}+o\left(\frac{1}{n^4}\right)\right)$$

$$=\frac{t}{n}+\frac{t^2}{n^3}+o\left(\frac{1}{n^3}\right)\to 0 \qquad (n\to\infty)$$

所以 $M_{S_n}(t)\to 1$. 由连续性定理 $S_n \xrightarrow{L} I$，其中I为退化分布 $\begin{pmatrix}0\\1\end{pmatrix}$ 随机变量.

4 二项分布的近似计算. 用正态分布对二项分布进行近似计算时，必须注意以下几点：

（1）当用积分极限定理去计算时，注意首先要把随机变量标准化后才能查正态表.

例4.2 若 $\zeta_n \sim b(500,\ 0.2)$，求 $P(100\leqslant\zeta_n\leqslant150)$.

注意 $P(100\leqslant\zeta_n\leqslant150)\neq\Phi(150)-\Phi(100)$，应先把 ζ_n 标准化，

而有

$$P(100\leqslant\zeta_n\leqslant150)=P\left(\frac{100-np}{\sqrt{npq}}\leqslant\frac{\zeta_n-np}{\sqrt{npq}}\leqslant\frac{150-np}{\sqrt{npq}}\right)$$

$$=\Phi\left(\frac{150-500\times0.2}{\sqrt{500\times0.2\times0.8}}\right)$$

$$-\Phi\left(\frac{100-500\times0.2}{\sqrt{500\times0.2\times0.8}}\right)$$

（2）不管是应用积分极限定理还是局部极限定理，还应注意 p 的大小，只有当 p 接近 $\frac{1}{2}$ 时，用此二定理做二项分布的近似计算结果才理想。当 p 很小(接近于 0 或 1 ）时，要考虑应用 *Poisson* 定理去进行近似计算。

（3）当 $\zeta_n\sim b(n,\ p)$时，求概率 $P(\zeta_n\leqslant b)$ 的近似值时，按理应求概率 $P(0\leqslant\zeta_n\leqslant b)$ 的近似值，而不应求 $P(-\infty<\zeta_n\leqslant b)$ 的近似值。但因

$$P(0\leqslant\zeta_n\leqslant b)=\Phi\left(\frac{b-np}{\sqrt{npq}}\right)-\Phi\left(\frac{-np}{\sqrt{npq}}\right)$$

而式中$\Phi\left(\frac{-np}{\sqrt{npq}}\right)$的值当 n 充分大时近于 0 ，这是因为

$$\frac{-np}{\sqrt{npq}}\leqslant\frac{-npq}{\sqrt{npq}}=-\sqrt{npq}\to-\infty$$

所以在应用积分极限定理去求二项分布随机变量 ζ_n 小于某个数 b 的概率的近似值时，实际上可以理解成求概率 $P(-\infty<\zeta_n\leqslant b)$ 的近似值，即只需计算$\Phi\left(\frac{b-np}{\sqrt{npq}}\right)$的值即可。

5　用频率估计概率时误差的估计．由积分极限定理，概率

$$P\left(\left|\frac{\mu_n}{n}-p\right|<\varepsilon\right)=P\left(-\varepsilon\sqrt{\frac{n}{pq}}<\frac{\mu_n-np}{\sqrt{npq}}<\varepsilon\sqrt{\frac{n}{pq}}\right)$$

$$\doteq\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right)-\Phi\left(-\varepsilon\sqrt{\frac{n}{pq}}\right)$$

$$=2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right)-1 \qquad (4-1)$$

这个关系可用来解决下列问题：

（1）已知 n, p, ε 时，求概率 $P\left(\left|\frac{\mu_n}{n}-p\right|<\varepsilon\right)$. 这时利用（4—1）式并查正态分布函数$\Phi(x)$的数值表可解决。

（2）要使$\frac{\mu_n}{n}$与 p 的差异不大于定数 ε 的概率不小于予先给定的数 β，问最少应做多少次试验，这时只需要求满足下式的最小 n，

$$2\Phi\left(\varepsilon\sqrt{\frac{n}{pq}}\right)-1\geqslant\beta$$

这也可通过查表求得。

（3）已知 n 及 β 求 ε，这类问题是在进行误差估计时提出来的。解法如下：先找 x_β，使

$$2\Phi(x_\beta)-1=\beta$$

这时 $\varepsilon=x_\beta\sqrt{\frac{pq}{n}}$即为所求。若 p 不知道，则利用 $pq\leqslant\frac{1}{4}$，有下列估计

$$\varepsilon\leqslant\frac{x_\beta}{2\sqrt{n}}$$

6　用矩母函数来讨论随机变量的性质

（1）用矩母函数计算随机变量的数字特征。有时我们通过随机变量的矩母函数来计算它的均值和方差更为方便。

例4.3 设ξ服从加马分布 $G(\alpha,\gamma)$，则 $M_\xi(t)=\left(\frac{\alpha}{\alpha-t}\right)^\gamma$ $(t<\alpha)$

因此　　$E\xi=M'_\xi(t)\Big|_{t=0}=\frac{\gamma}{\alpha}$

而　　$E\xi^2=M''_\xi(t)\Big|_{t=0}=\gamma(\gamma+1)/\alpha^2$

所以　　$D\xi=\frac{\gamma(\gamma+1)}{\alpha^2}-\left(\frac{\gamma}{\alpha}\right)^2=\frac{\gamma}{\alpha^2}$

令 $\alpha=\frac{1}{2}$，$\gamma=\frac{n}{2}$，则 ξ 为$\chi^2(n)$分布。这时$E\xi=n$，$D\xi=2n$。

例 4.4 设 $\xi \sim N(\mu,\sigma^2)$，则 $M_\xi(t)=exp\left\{\mu t+\frac{\sigma^2 t^2}{2}\right\}$

于是 $$E\xi=M'_\xi(t)\Big|_{t=0}=(\mu+\sigma^2 t)M_\xi(t)\Big|_{t=0}=\mu$$

而 $$E\xi^2=M''_\xi(t)\Big|_{t=0}=[M_\xi(t)\sigma^2+(\mu+\sigma^2 t)^2 M_\xi(t)]_{t=0}$$
$$=\sigma^2+\mu^2$$

这样 $D\xi=\sigma^2$。

（2）用矩母函数讨论独立随机变量和的分布．矩母函数是研究独立随机变量和的分布（加法定理）的有效工具．

例 4.5 设 $\xi_i(i=1,2,\cdots,k)$相互独立，$\xi_i\sim b(n_i,p)$则 $\zeta_k=\sum\limits_{i=1}^{k}\xi_i\sim b\left(\sum\limits_{i=1}^{k}n_i,\ p\right)$。

事实上，ξ_i 的矩母函数 $M_{\xi_i}(t)=(q+pe^t)^{n_i}$，故 ζ_k 的矩母函数

$$M_{\zeta_k}(t)=\prod_{i=1}^{k}M_{\xi_i}(t)=\prod_{i=1}^{k}(q+pe^t)^{n_i}$$
$$=(q+pe^t)^{n_1+n_2+\cdots+n_k}$$

这正是参数为 $n=n_1+n_2+\cdots+n_k$ 和 p 的二项分布的矩母函数，

由唯一性定理，$\zeta_k\sim b\left(\sum\limits_{i=1}^{k}n_i,\ p\right)$。

例 4.6 设 $\xi_i(i=1,2,\cdots,n)$相互独立服从参数为 λ_i 的普阿松分布．则 $\zeta_n=\sum\limits_{i=1}^{n}\xi_i$ 仍然服从普阿松分布，其参数为 $\lambda=\sum\limits_{i=1}^{n}\lambda_i$。

事实上，ξ_i 的矩母函数为 $M_{\xi_i}(t)=exp\{\lambda(e^t-1)\}$，所以 ζ_n 的矩母函数

$$M_{\zeta_n}(t)=\prod_{i=1}^{n}M_{\zeta_i}(t)=\prod_{i=1}^{n}exp\{\lambda_i(e^t-1)\}$$

$$=\exp\left\{(e^t-1)\sum_{i=1}^{n}\lambda_i\right\}$$

可见 $\zeta_n \sim P\left(\sum_{i=1}^{n}\lambda_i\right)$.

例 4.7 若 $\xi_i \sim G(\alpha, \gamma_i)$ $(i=1,2,\cdots,n)$，且相互独立，则 $\zeta_n=\sum_{i=1}^{n}\xi_i \sim G\left(\alpha, \sum_{i=1}^{n}\gamma_i\right)$.

因 $M_{\xi_i}(t)=\left(\frac{\alpha}{\alpha-t}\right)^{\gamma_i}$，所以

$$M_{\zeta_n}(t)=\prod_{i=1}^{n}\left(\frac{\alpha}{\alpha-t}\right)^{\gamma_i}=\left(\frac{\alpha}{\alpha-t}\right)^{\sum_{i=1}^{n}\gamma_i}$$

例 4.8 若 $\xi_i \sim N(\mu_i, \sigma_i^2)$ $(i=1,2,\cdots,n)$，且相互独立，则 $\zeta_n=\sum_{i=1}^{n}\xi_i \sim N\left(\sum_{i=1}^{n}\mu_i, \sum_{i=1}^{n}\sigma_i^2\right)$.

事实上，$M_{\xi_i}(t)=\exp\left\{\mu_i t+\frac{\sigma_i^2 t^2}{2}\right\}$，

所以

$$M_{\zeta_n}(t)=\prod_{i=1}^{n}\exp\left\{\mu_i t+\frac{\sigma_i^2 t^2}{2}\right\}$$

$$=\exp\left\{t\sum_{i=1}^{n}\mu_i+\frac{t^2}{2}\sum_{i=1}^{n}\sigma_i^2\right\}$$

（3）用矩母函数证明随机变量的独立性。

我们可以证明，对任意随机变量 ξ，η，它们是相互独立的充要条件是对一切实数 t_1，t_2 有

$$Ee^{t_1\xi+t_2\eta}=M_\xi(t_1)M_\eta(t_2).$$

事实上，若 ξ，η 的密度函数及联合分布密度分别为 $f_\xi(x)$，$f_\eta(y)$ 和 $f(x,y)$。由 ξ 与 η 的独立性，立刻推出

$$Ee^{t_1\xi+t_2\eta}=Ee^{t_1\xi}\cdot Ee^{t_2\eta}=M_\xi(t_1)\cdot M_\eta(t_2).$$

反之，若 $Ee^{t_1\xi+t_2\eta}=M_\xi(t_1)\cdot M_\eta(t_2)$，则有

$$\iint e^{t_1x+t_2y}f(x,y)dxdy=\left[\int e^{t_1x}f_\xi(x)dx\right]\left[\int e^{t_2y}f_\eta(y)dy\right]$$
$$=\iint e^{t_1x+t_2y}f_\xi(x)f_\eta(y)dxdy$$

由矩母函数的唯一性定理*，有

$$f(x,\ y)=f_\xi(x)f_\eta(y)$$

这说明 ξ，η 相互独立。

例 4.9　若 $\xi_i(i=1,2,\cdots,n)$ 相互独立，$\xi_i\sim N(\mu_i,\ \sigma_i^2)$，$a_1$，$a_2,\cdots,a_n$；$b_1$，$b_2$，$\cdots,b_n$ 为非零实数，满足 $\sum\limits_{i=1}^{n}a_ib_i\sigma_i^2=0$，则 $L_1=\sum\limits_{i=1}^{n}a_i\xi_i$ 和 $L_2=\sum\limits_{i=1}^{n}b_i\xi_i$ 相互独立。

证明　不失普遍性，令 $\mu_i=0(i=1,2,\cdots,n)$。对任意实数 α，β 和 t，

$$Ee^{(\alpha L_1+\beta L_2)t}=E\exp\left\{t\sum_{i=1}^{n}(\alpha a_i+\beta b_i)\xi_i\right\}$$
$$=\prod_{i=1}^{n}\exp\left\{\frac{t^2}{2}(\alpha a_i+\beta b_i)^2\sigma_i^2\right\}$$
$$=\exp\left\{\frac{\alpha^2t^2}{2}\sum_1^n a_i^2\sigma_i^2\right.$$
$$\left.+\frac{\beta^2t^2}{2}\sum_1^n b_i^2\sigma_i^2\right\}\quad\left(\text{因}\sum_1^n a_ib_i\sigma_i^2=0\right)$$
$$=\prod_{i=1}^{n}\exp\left\{\frac{t^2\alpha^2}{2}a_i^2\sigma_i^2\right\}\prod_{i=1}^{n}\exp\left\{\frac{t^2\beta^2}{2}b_i^2\sigma_i^2\right\}$$
$$=\prod_{i=1}^{n}Ee^{t\alpha a_i\xi_i}\prod_{i=1}^{n}Ee^{t\beta b_i\xi_i}$$

---

* 二维随机变量 (ξ,η) 的矩母函数可类似于一维情况定义为二元函数 $M(t_1,t_2)=Ee^{t_1\xi+t_2\eta}$。对二维场合的矩母函数也有类似一维的唯一性定理。

$$= Eexp\left\{t\alpha\sum_1^n a_i\xi_i\right\}\cdot Eexp\left\{t\beta\sum_1^n b_i\xi_i\right\}$$

$$= Ee^{\alpha t\zeta_1}\cdot Ee^{\beta t\zeta_2}$$

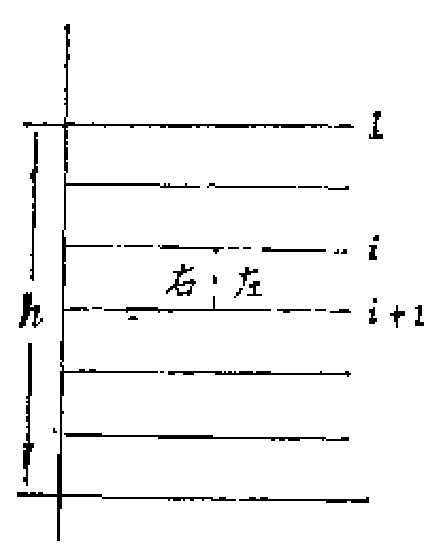
作为特例，当ξ，η相互独立时，$\xi\sim N(\mu_1,\sigma^2)$，$\eta\sim N(\mu_2,\ \sigma^2)$ 则 $\xi-\eta$，$\xi+\eta$是相互独立的。

7　中心极限定理的一个直观例子

现在我们来考虑这样一个问题，为什么农民在脱谷‘扬场’时，最后扬出的是一个‘钟形’谷堆呢？为了解释这个现象，考虑一个简化了的模型，即在本章习题6所讨论的在钉板里小球下落的问题．通过此题的计算我们知道，小球在每一层上落入下一层时只有两种情况，不是向左就是向右，落入第n层的位置记为ξ_n，则$\xi_n\sim b\left(n-1,\ \frac{1}{2}\right)$，即

$$P(\xi_n=k)=\binom{n-1}{k}\left(\frac{1}{2}\right)^{n-1},\ k=0,1,\cdots,n-1$$

下面我们转回来考虑在‘扬场’时米粒下落的问题。设米粒扬起的高度为h，我们设想把h分割成n个等分，假想米粒是从最高点一层一层地下落。现在考虑当米粒处于第i层正向第$i+1$层降落的情形，我们假想从它在第i层所处的位置向下作一条垂线，这条垂线就把第$i+1$层分为左、右两部分，由于米粒是随机下落的，所以它不是落在这条垂线的左边就是落在它的右边，如图(4-1)，这就正像上面开始时考虑的球在钉板里从第i层上继续下落的情形。因此米粒在第n层上所处的位置$\xi_n\sim b\left(n-1,\ \frac{1}{2}\right)$．现在如果把所分的层加细，即令$n\to\infty$，则上述模型就完全刻划了米粒下落



图（4—1）

的实际状况。由积分极限定理，二项分布渐近于正态分布。所以米粒最后堆积的曲线便是中间高两边低呈对称的钟形曲线了。

这个例子实际上给出的是一个n个相互独立同两点分布随机变量和渐近正态分布的直观实例。

习题解答

1　(a)由矩母函数的定义

$$M_\xi(t)=\int_{-\infty}^{\infty}e^{tx}f(x)dx$$

$$=\int_0^1 e^{tx}2xdx$$

设　$e^{tx}dx=dv,\ u=x\ \Rightarrow\ v=\frac{1}{t}e^{tx},du=dx$

用分部积分法，得

$$M_\xi(t)=2\left(\frac{1}{t}e^{tx}\cdot x\Big|_0^1-\frac{1}{t}\int_0^1 e^{tx}dx\right)$$

$$=2\left(\frac{1}{t}e^t-\frac{1}{t^2}e^t+\frac{1}{t^2}\right)=\frac{2(te^t-e^t+1)}{t^2}$$

(b)因为

$$M'_\xi(t)=2\left(\frac{e^t}{t}-2\frac{e^t}{t^2}+\frac{2(e^t-1)}{t^3}\right)$$

$$=2\frac{t^2e^t-2te^t+2(e^t-1)}{t^3}$$

$$M''_\xi(t)=2\left(\frac{te^t-e^t}{t^2}-2\frac{t^2e^t-2te^t}{t^4}\right.$$

$$\left.+\frac{2t^3e^t-6t^2(e^t-1)}{t^6}\right)$$

$$=2\frac{t^3e^t-3t^2e^t+6te^t-6(e^t-1)}{t^4}$$

所以　$E\xi=M'_\xi(0)=\lim\limits_{t\to 0}M'_\xi(t)$

$$=\lim_{t\to 0}2\frac{2te^{t}+t^{2}e^{t}-2e^{t}-2te^{t}+2e^{t}}{3t^{2}}$$

$$=\lim_{t\to 0}\frac{2e^{t}}{3}=\frac{2}{3}$$

$$E\xi^{2}=M_{\xi}''(0)$$

$$=\lim_{t\to 0}M''_{\xi}(t)$$

$$=\lim_{t\to 0}2\cdot\frac{3t^{2}e^{t}+t^{3}e^{t}-6te^{t}-3t^{2}e^{t}+6e^{t}+6te^{t}-6e^{t}}{4t^{3}}$$

$$=\lim_{t\to 0}\frac{2e^{t}}{4}=\frac{1}{2}$$

$$D\xi=E\xi^{2}-(E\xi)^{2}$$

$$=\frac{1}{2}-\left(\frac{2}{3}\right)^{2}$$

$$=\frac{1}{18}$$

2 (a) $M_{\xi}(t)=\int_{-\infty}^{\infty}e^{tx}f(x)dx$

$$=\int_{a}^{\infty}e^{tx}\lambda e^{-\lambda(x-a)}dx$$

$$=\lambda\int_{a}^{\infty}e^{tx-\lambda(x-a)}\,dx=\frac{\lambda}{t-\lambda}\,e^{tx-\lambda(x-a)}\Big|_{a}^{\infty}$$

$$=\frac{\lambda}{t-\lambda}e^{(t-\lambda)x+\lambda a}\Big|_{a}^{\infty}=\frac{\lambda}{\lambda-t}e^{ta}\qquad(t<\lambda\text{时})$$

(b)先计算$M'_{\xi}(t)$，$M''_{\xi}(t)$。

$$M'_{\xi}(t)=\frac{\lambda(ae^{ta}(\lambda-t)+e^{ta})}{(\lambda-t)^{2}}$$

$$M''_{\xi}(t)=\lambda\left[\frac{\{a^{2}e^{ta}(\lambda-t)-ae^{ta}+ae^{ta}\}(\lambda-t)^{2}}{(\lambda-t)^{4}}+\frac{2(\lambda-t)\{ae^{ta}(\lambda-t)+e^{ta}\}}{(\lambda-t)^{4}}\right]$$

$$=\lambda(\lambda-t)\left[\frac{a^2e^{ta}(\lambda-t)(\lambda-t)}{(\lambda-t)^4}+\frac{2(ae^{ta}(\lambda-t)+e^{ta})}{(\lambda-t)^4}\right]$$

$$E\xi=M'_\xi(0)=\frac{\lambda(a\lambda+1)}{\lambda^2}=\frac{a\lambda+1}{\lambda}$$

$$E\xi^2=M''_\xi(0)=\frac{\lambda^2[a^2\lambda^2+2(a\lambda+1)]}{\lambda^4}$$

$$=\frac{a^2\lambda^2+2a\lambda+2}{\lambda^2}$$

所以

$$D\xi=\frac{a^2\lambda^2+2a\lambda+2}{\lambda^2}-\left(\frac{a\lambda+1}{\lambda}\right)^2$$

$$=\frac{1}{\lambda^2}$$

3

因 $\xi\sim N(\mu_\xi,\ \sigma_\xi^2)$

$\eta\sim N(\mu_\eta,\ \sigma_\eta^2)$

故它们的矩母函数分别为

$$M_\xi(t)=exp\left(\mu_\xi t+\frac{1}{2}\sigma_\xi^2t^2\right)$$

$$M_\eta(t)=exp\left(\mu_\eta t+\frac{1}{2}\sigma_\eta^2t^2\right)$$

而 $a\xi$，$b\eta$ 的矩母函数分别为

$$M_{a\xi}(t)=M_\xi(at)=exp\left(\mu_\xi at+\frac{1}{2}\sigma_\xi^2a^2t^2\right)$$

$$M_{b\eta}(t)=M_\eta(bt)=exp\left(\mu_\eta bt+\frac{1}{2}\sigma_\eta^2b^2t^2\right)$$

因 ξ，η 相互独立，所以 $a\xi$ 与 $b\eta$ 相互独立，故 $\zeta=a\xi+b\eta$的矩母函数为

$$M_\zeta(t)=M_\xi(at)\cdot M_\eta(bt)$$

$$=exp\left\{(\mu_\xi a+\mu_\eta b)t+\frac{1}{2}(\sigma_\xi^2a^2+\sigma_\eta^2b^2)t^2\right\}$$

这正是具有均值为 $\mu_\xi a+\mu_\eta b$，方差为 $\sigma_\xi^2a^2+\sigma_\eta^2b^2$ 的正态随机变量的矩母函数，可见

$$\zeta\sim N(\mu_\xi a+\mu_\eta b,\ \sigma_\xi^2a^2+\sigma_\eta^2b^2)$$

4　(a) $\eta=3\xi+2$ 的矩母函数

$$M_\eta(t)=e^{2t}M_\xi(3t)$$
$$=e^{2t}(0.4e^{3t}+0.6)^8$$

(b) $M_\xi'(t)=8(0.4e^t+0.6)^7 0.4e^t$

$$E\xi=M_\xi'(0)=3.2$$

5　$M_\xi(t)=M_{\xi_1}(t)\cdot M_{\xi_2}(t)=exp\{\lambda_1(e^t-1)\}\cdot exp\{\lambda_2(e^t-1)\}$
$$=exp\{(\lambda_1+\lambda_2)(e^t-1)\}$$

故 ξ 服从参数为 $\lambda_1+\lambda_2$ 的 *Poisson* 分布。

6　球落到第 n 层，它有 n 个可能位置，设这 n 个位置从左至右分别为 0，1，…，$n-1$。下面我们求球落在第 $k(0\leqslant k\leqslant n-1)$ 个位置的概率。因为在每一层各个可能点上的球，如继续下落都有两个可能落法：不是向左就是向右。向左或向右的概率均为 $\frac{1}{2}$。球下落到第 n 层恰进行了 $n-1$ 次下落。这样显然，球落到第 n 层第 k 个位置，当且仅当球在 $n-1$ 次下落时恰有 k 次向右而 $n-1-k$ 次向左。如果把球下落一层看成一次试验，则这种试验只有两个可能结果（向左，或向右），每次相互独立，下落到 n 层就相当于独立进行了 $n-1$ 次这种试验。因此若第 n 层上球落入的位置设为 ξ_n，则 ξ_n 服从参数为 $\left(n-1,\frac{1}{2}\right)$ 的二项分布。即

$$P(\xi_n=k)=\binom{n-1}{k}\left(\frac{1}{2}\right)^{n-1}\qquad k=0,1,\cdots,n-1$$

当球不断放入时，球的分布将呈现中间高，两边低，成对称的样子，由积分极限定理知它渐近于正态分布。

7　设 ξ 为400粒种子中不发芽的种子数，显然 $\xi\sim b(400,0.05)$。

因此

$$P(0\leqslant\zeta\leqslant 25)=P\left(\frac{0-400\times 0.05}{\sqrt{400\times 0.05\times 0.95}}\leqslant\frac{\zeta-np}{\sqrt{npq}}\right.$$

$$\left.\leqslant\frac{25-400\times 0.05}{\sqrt{400\times 0.05\times 0.95}}\right)$$

$$=P\left(-4.588\leqslant\frac{\zeta-np}{\sqrt{npq}}\leqslant 1.147\right)$$

$$\doteq\frac{1}{\sqrt{2\pi}}\int_{-4.588}^{1.147}e^{-\frac{x^2}{2}}dx$$

$$=\Phi(1.147)-\Phi(-4.588)$$

$$=\Phi(1.147)-1+\Phi(4.588)$$

$$=0.8749-1+0.999997\doteq 0.8749$$

8

(a) $\zeta=\xi_1+\xi_2+\cdots+\xi_{50}$

$$E\zeta=\sum_{i=1}^{50}E\xi_i=50\lambda=50\times 0.03-1.5$$

$$D\zeta=\sum_{i=1}^{50}D\xi_i=50\lambda=1.5$$

$$P(\zeta\geqslant 3)=P\left(\frac{3-1.5}{\sqrt{1.5}}\leqslant\frac{\zeta-E\zeta}{\sqrt{D\zeta}}\right)$$

$$=P\left(1.23\leqslant\frac{\zeta-E\zeta}{\sqrt{D\zeta}}\right)$$

$$\doteq 1-\Phi(1.23)=1-0.8907=0.1093.$$

(b) ζ服从参数为 $\lambda=1.5$ 的 *Poisson* 分布

$$P(\zeta\geqslant 3)=\sum_{k=3}^{\infty}\frac{1.5^k}{k!}e^{-1.5}\doteq 0.1912$$

9　显然 $P=\frac{3}{4}$，ζ 服从参数为 $\left(n,\ \frac{3}{4}\right)$的二项分布，而

$$0.9964=P\left(\left|\frac{\zeta}{n}-P\right|\leqslant 0.001\right)$$

$$= P\left(-0.001 \leqslant \frac{\zeta}{n} - p \leqslant 0.001\right)$$

$$= P\left(-0.001\sqrt{\frac{n}{pq}} \leqslant \frac{\zeta - np}{\sqrt{npq}} \leqslant 0.001\sqrt{\frac{n}{pq}}\right)$$

$$\doteq \Phi\left(0.001\sqrt{\frac{n}{pq}}\right) - \Phi\left(-0.001\sqrt{\frac{n}{pq}}\right)$$

$$= 2\Phi\left(0.001\sqrt{\frac{n}{pq}}\right) - 1$$

即 $$\Phi\left(0.001\sqrt{\frac{n}{pq}}\right) \doteq \frac{1.9964}{2} = 0.9982$$

由正态表查得$0.001\sqrt{\frac{n}{pq}} = 2.92$

于是 $n = \left(\frac{\sqrt{pq} \times 2.92}{0.001}\right)^2 = \frac{\frac{3}{4} \cdot \frac{1}{4} \times 2.92^2}{0.001^2} = 1598700$

10 ξ_i的分布密度 $p(x) = \begin{cases} \frac{1}{T} & x \in [0,T] \\ 0 & \text{其 他} \end{cases}$

则随机变量 $\zeta_2 = \xi_1 + \xi_2$ 的分布密度

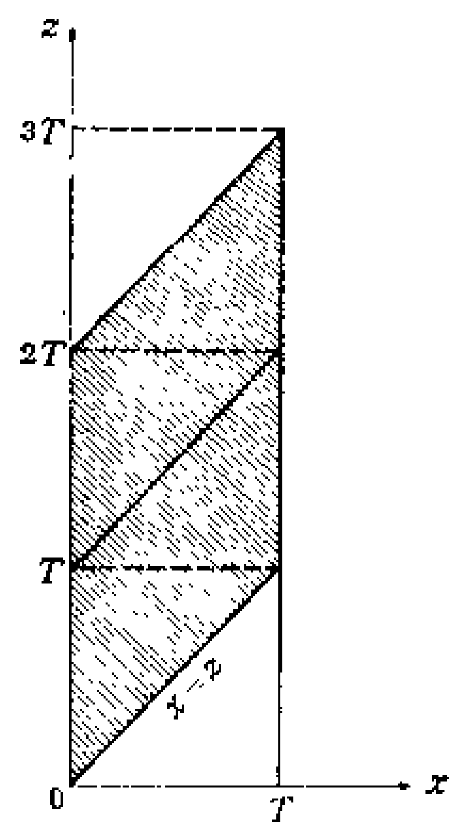
$$P_{\zeta_2}(z) = \int_{-\infty}^{\infty} p(x)p(z-x)dx$$

$$= \begin{cases} z/T^2 & 0 \leqslant z \leqslant T \\ (2T - z)/T^2 & T \leqslant z \leqslant 2T \\ 0 & \text{其 他} \end{cases}$$

$\zeta_3 = \xi_1 + \xi_2 + \xi_3 = \zeta_2 + \xi_3$ 的分布密度

$$P_{\zeta_3}(z) = \int_{-\infty}^{\infty} p(x)p_{\zeta_2}(z-x)dx$$

被积函数的取值非零区域为 $0 \leqslant x \leqslant T$，$0 \leqslant z - x \leqslant 2T$， 如图(4—10题)中的阴影部分．当$0 \leqslant z \leqslant T$时，



(图 4—10题)

$$p_{\zeta_3}(z)=\int_0^z \frac{1}{T^3}(z-x)dx=\frac{z^2}{2T^3}$$

当 $T\leqslant z\leqslant 2T$ 时，

$$p_{\zeta_3}(z)=\int_0^{z-T}\frac{1}{T^3}(2T-z+x)dx+\int_{z-T}^{T}\frac{1}{T^3}(z-x)dx$$

$$=\frac{1}{2T^3}(-2z^2+6Tz-3T^2)$$

当 $2T\leqslant z\leqslant 3T$ 时，

$$p_{\zeta_3}(z)=\int_{z-2T}^{T}\frac{1}{T^3}(2T-z+x)dx$$

$$=\frac{1}{2T^3}(3T-z)^2$$

容易计算

$$E\xi_i=\frac{T}{2},\qquad D\xi_i=\frac{T^2}{12}$$

$$E\zeta_2=T,\qquad D\zeta_2=\frac{T^2}{6}$$

$$E\zeta_3=\frac{3T}{2},\qquad D\zeta_3=\frac{T^2}{4}$$

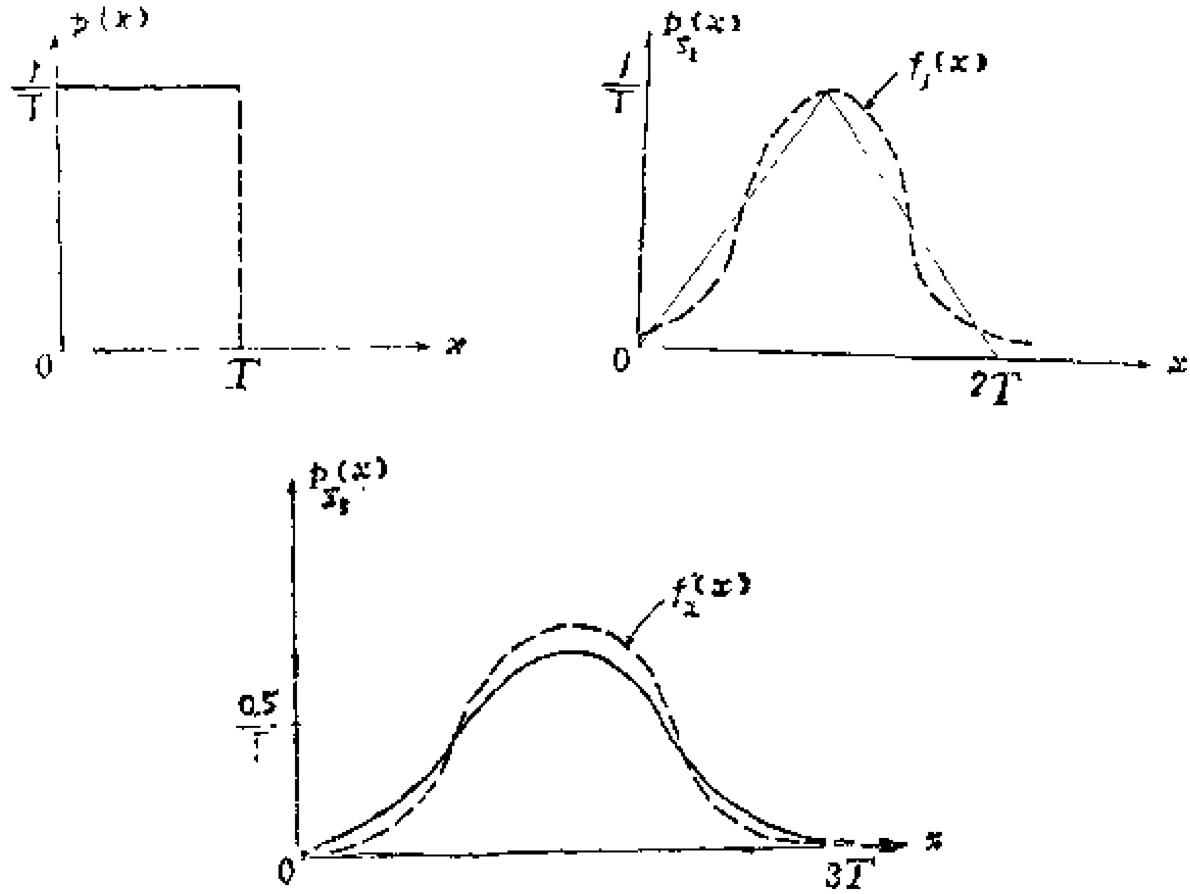
从 ζ_2 的密度函数 $P_{\zeta_2}(z)$ 的表达式可以看出，它是一个三角形分布，从 $P_{\zeta_3}(z)$ 的表达式可以看出 ζ_3是一个由三个抛物线段所组成的一条分布曲线．与ζ_2 及ζ_3 的均值和方差相同的正态密度分别为

$$f_1(x)=\frac{1}{T}\sqrt{\frac{3}{\pi}}e^{-3(x-T)^2/T^2}$$

和

$$f_2(x)=\frac{1}{T}\sqrt{\frac{2}{\pi}}e^{-2(x-1.5T)^2/T^2}$$

各分布密度曲线如图（4—10题）



（图 4—10题）

b

11　记$U_n=\frac{1}{\sqrt{n}}(\xi_1+\xi_2+\cdots+\xi_n)$，$V_n=\frac{\xi_1^2+\xi_2^2+\cdots+\xi_n^2}{n}$则 $W_n=\frac{U_n}{V_n}$。U_n 的矩母函数

$$M_{V_n}(t)=\prod_{i=1}^{n}Ee^{t\xi_i/\sqrt{n}}=\prod_{i=1}^{n}e^{t^2/2n}=e^{t^2/2}$$

因此 U_n是正态 $N(0,1)$随机变量，故$U_n\xrightarrow{L}Z$，其中 Z 是 $N(0,1)$随机变量．对于V_n，我们注意到每一个 ξ_i^2 都是 $\chi^2(1)$，所以 V_n 的矩母函数

$$M_{V_n}(t)=\prod_{i=1}^{n}\left(\frac{1}{1-2t/n}\right)^{\frac{1}{2}}\qquad t<\frac{n}{2}$$

$$=\left(1-\frac{2t}{n}\right)^{-n/2}\qquad t<\frac{n}{2}$$

它正是具有参数为 $\gamma=n/2$, $\alpha=n/2$ 的 *Gamma* 分布的矩母函数. $M_{V_n}(t)\to e^t$ 当 $n\to\infty$ 时. 所以 $V_n\xrightarrow{P}1$, 由定理 4.1.8 $W_n=U_n/V_n\xrightarrow{L}Z$.

12 ξ_n 的矩母函数 $M_n(t)=(1/n)e^t+[1-(1/n)]$ 对一切 $t\in R$ 存在且, 当 $n\to\infty$ 时,对一切 t 成立 $M_n(t)\to 1$. 因为 1 是退化随机变量 $\xi:\begin{pmatrix}0\\1\end{pmatrix}$ 的矩母函数, 所以得证 $\xi_n\xrightarrow{L}\xi$.

13 ξ_λ 的矩母函数$M(t)=exp\{\lambda(e^t-1)\}$,令 $\eta=(\xi_\lambda-\lambda)/\sqrt{\lambda}$, 则 η 的矩母函数

$$M_\eta(t)=e^{-t\sqrt{\lambda}}M\left(\frac{t}{\sqrt{\lambda}}\right)$$

取对数 $\ln M_\eta(t)=-t\sqrt{\lambda}+\ln M\left(\frac{t}{\sqrt{\lambda}}\right)$

$$=-t\sqrt{\lambda}+\lambda(e^{t/\sqrt{\lambda}}-1)$$

$$=-t\sqrt{\lambda}+\lambda\left(\frac{t}{\sqrt{\lambda}}+\frac{t^2}{2\lambda}+\frac{t^3}{3!\lambda^{3/2}}+\cdots\right)$$

$$=\frac{t^2}{2}+\frac{t^3}{3!\lambda^{1/2}}+\cdots$$

当$\lambda\to\infty$ 时,有$\ln M_\eta(t)\to\frac{t^2}{2}$, 所以 $M_\eta(t)\to e^{t^2/2}$, 即 $\eta\xrightarrow{L}\xi$, ξ 为$N(0,1)$随机变量.

14 由切贝晓夫不等式

$$P\left(\left|\sum_{i=1}^n\xi_i-\sum_{i=1}^n\mu_i\right|\geq\varepsilon\sum_{i=1}^n\sigma_i^2\right)\leqslant\frac{E\left(\sum_{i=1}^n(\xi_i-\mu_i)\right)^2}{\varepsilon^2\left(\sum_{i=1}^n\sigma_i^2\right)^2}$$

$$=\frac{1}{\varepsilon^2\sum_{i=1}^n\sigma_i^2}\to 0\quad 当n\to\infty时.$$

第五章 统 计 推 断

前几章讲了概率论的基本知识。在那里我们研究随机变量时，总是假定所研究的随机变量的概率分布或某些数字特征为已知，以此为前提进行理论讨论。比如，已知随机变量X的概率分布求X落在某个范围里的概率或求X的均值，方差；已知某些随机变量的概率分布或数字特征求另外与之有关的随机变量的概率分布或数字特征，等等。可是，在很多实际问题中，我们对所考虑的随机变量，这些都是不知道的。比如，灯泡厂每天生产上万个灯泡，这些灯泡的次品率p是多少？灯泡寿命X的概率分布怎么样？这又涉及到寿命的均值EX及寿命的方差DX等于什么？这些都不知道。但是，我们又必须及时掌握每天生产的大批灯泡的质量，可行的办法就只能是通过抽样了解其中的一部分，比如10个，20个的质量情况，然后根据这一部分的质量情况对整批灯泡的质量进行估计或判断。从这个例子可以看出，对所研究的随机现象，如何取得有用的观察资料，如何根据取得的有限资料，对所研究的随机现象的总体的统计规律性进行科学的分析，尽量作出精确可靠的推断，就是非常重要、非常实际的问题。数理统计就是研究这一类问题的数学分支。在数理统计学中，由于只能依据有限资料去推断总体的规律，因而作出的结论不可能绝对准确，总带有一定程度的不确定性。我们用概率来刻划这种不确定性。概率大，推断较可靠；概率小，推断较不可靠。这种伴随有一定概率的推断，在数理统计学中就称为统计推断。

在这一章里，我们讨论以下三个方面的问题：

一 点估计，

二　显著性检验；

三　区间估计。

§5.1 总体与子样

(一) 基本概念

首先，我们引入一些数理统计学中常用的概念与术语。

在日常生活中，我们把所要考察的对象的全体称为总体，而把组成总体的每一个成员称为个体。例如，在考察某批灯泡的质量时，该批灯泡的全体就是一个总体，其中的每一只灯泡就是个体。又如，在考察某个品种水稻的性状时，这个品种的植株的全体就是一个总体，而每一棵水稻植株就是个体。

但是，在数理统计学中，我们并不笼统地研究所关心的对象，而只是对它的数值指标感到兴趣。例如，考察灯泡时，我们并不研究它的形状、式样等特征，而只是关心灯泡寿命、灯泡亮度等数值指标的大小。当我们只考察灯泡寿命这项指标时，一批灯泡中的每一只灯泡均有一个确定的寿命值，因此，很自然地，我们应该把所有这些寿命值的全体当成总体，这时，每个灯泡的寿命值就是个体。

我们知道，既使在相同的生产条件下生产的灯泡，由于种种微小的偶然因素的影响，它们的寿命值也不尽相同，但却服从确定的统计规律。这表明灯泡寿命X是一个随机变量。这时，每只灯泡的寿命值就是它的可能值，而总体就是它的所有这些可能值的全体。当批量很大时，我们干脆就用X来代表这一总体，而X的可能值就是它的个体。如果X具有分布函数$F(x)$，则称这一总体为具有分布函数$F(x)$的总体。

在一个总体（例如一批灯泡，考虑其寿命）X中，抽取n个个体X_1，X_2，$\cdots$，X_n（实际上，相当于对灯泡寿命作了n次观测，

得到 n 个个体的寿命值），这 n 个个体 X_1、X_2，…，X_n 称为总体 X 的一个容量为 n 的子样（或样本）。子样容量 n 也可称为子样的大小。

应该注意的是，子样 $(X_1, X_2, \cdots, X_n)$ 是 n 维随机向量。一批抽样所得到的 n 个具体数值 $(x_1, x_2, \cdots, x_n)$ 只是这个 n 维随机向量的一个可能值，称为子样 $(X_1, X_2, \cdots, X_n)$ 的观测值或样本值。所有样本值的全体构成样本空间，以后称为子样空间。一个子样观测值 $(x_1, x_2, \cdots, x_n)$ 就是子样空间中的一个点。

抽取子样的目的在于对总体的统计规律进行推断。因而，很自然地要求所抽取的子样能够很好地反映总体的特性。由此，必须对抽样方法有所要求，通常提出下面两条：（1）代表性，X_i 的分布与总体 X 的分布 $F(x)$ 相同；（2）独立性，$X_1, X_2, \cdots, X_n$ 为相互独立的随机变量。满足这两条要求的子样称为简单随机子样。本书只讨论简单随机子样。以后，如无特殊的声明，所说的子样都是指的简单随机子样。

总起来讲，总体就是一个随机变量 X。子样就是一组 n 个相互独立的与 X 有相同概率分布的随机变量 $X_1, X_2, \cdots, X_n$，也可看成一个随机向量 $(X_1, X_2, \cdots, X_n)$。每一次抽样时，我们得到的一组 n 个具体的数值 $(x_1, x_2, \cdots, x_n)$，是这一随机向量的一个可能值（样本值）。

对于简单随机子样 $(X_1, X_2, \cdots, X_n)$，其分布由总体 X 的分布函数 $F(x)$ 完全确定。特别，当 X 为连续型随机变量，其分布密度为 $p(x)$ 时，$(X_1, X_2, \cdots, X_n)$ 的联合分布密度为 $p(x_1)\cdot p(x_2)\cdots p(x_n)$。

例5.1.1 一批产品共 N 件，废品率为 p。以 X 记数值指标，废品为1，正品为0。我们有放回地重复取样，子样容量为 n，则子样 $(X_1, X_2, \cdots, X_n)$ 就是简单随机子样。每个 X_i 均只能取0，1两个数值。因此子样空间由 2^n 个点 $(x_1, x_2, \cdots, x_n)$ 组成，x_i 取0或1，$1 \leqslant i \leqslant n$。实际上，抽样时是不会采用有放回取样的。通

常用无放回取样。这时，严格地讲，子样不是简单随机子样，因此诸 X_i 的独立性不能成立。特别，当 $n=2$ 时，

$$P(X_1=1, X_2=1)=P(X_1=1)\cdot P(X_2=1|X_1=1)$$
$$=\frac{Np}{N}\cdot\frac{Np-1}{N-1}$$
$$\neq p\cdot p=P(X_1=1)\cdot P(X_2=1)。$$

但是，当 N 足够大时，$\frac{Np-1}{N-1}$ 与 p 相差无几，这样就可以近似地看成简单随机子样。

以后，我们还将经常用到子样的一些函数。称

$$\overline{X}\triangleq\frac{1}{n}(X_1+X_2+\cdots+X_n)$$

为子样均值；

$$S^2\triangleq\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

为子样方差；

$$\hat{m}_r\triangleq\frac{1}{n}\sum_{i=1}^{n}X_i^r$$

为子样的 r 阶矩（或 r 阶原点矩）；

$$\hat{\mu}_r\triangleq\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^r$$

为子样的 r 阶中心矩。

（二）正态总体下，$\overline{X}$ 与 S^2 的分布

由于子样均值 $\overline{X}$ 与子样方差 S^2 在数理统计中有广泛的应用，应该对它们详加讨论。下面，我们推导，当总体 X 服从正态分布时，$\overline{X}$ 与 S^2 的分布。

这一段需要一点线性代数的基础知识，如果读者不具有这样的

准备，则可以跳过不读，只要会用有关的结论即可。

以下，矩阵或向量均用大写的粗体拉丁字母表示，如 $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C},\cdots$；$\boldsymbol{X}$，$\boldsymbol{Y}$，$\boldsymbol{Z}$，$\cdots$。

I_n 记 n 阶单位阵。1_n 记元素全为 1 的 n 维列向量。有时脚标也可略去不写，这时阶数（维数）视上下文而定。

$\boldsymbol{A}'$ 记$\boldsymbol{A}$的转置阵，$rk\boldsymbol{A}$ 记 $\boldsymbol{A}$的秩，$tr\boldsymbol{A}$ 记$\boldsymbol{A}$的迹，即 $tr\boldsymbol{A}=\sum_{i=1}^{n}a_{ii}$ 为$\boldsymbol{A}$的所有主对角线上元素之和。

若$\boldsymbol{X}=(X_{ij})_{n\times p}$，其中 X_{ij} 均为随机变量，则称$\boldsymbol{X}$为$n\times p$阶随机矩阵。定义矩阵$(EX_{ij})_{n\times p}$ 为$\boldsymbol{X}$的均值矩阵，并记为$E\boldsymbol{X}$，即

$$E\boldsymbol{X}=(EX_{ij})_{n\times p}$$

特别地，当 $n=1$ 时，$\boldsymbol{X}=(X_{11}, X_{12}, \cdots, X_{1p})$是 p 维随机（行）向量，这时 $E\boldsymbol{X}=(EX_{11}, EX_{12}, \cdots, EX_{1p})$。当 $p=1$ 时，$\boldsymbol{X}$ 与 $E\boldsymbol{X}$均是n 维（列）向量。

引理 1　若$\boldsymbol{X}$是$n\times p$阶随机矩阵，$\boldsymbol{A}=(a_{li})_{m\times n}$，$\boldsymbol{B}=(b_{jk})_{p\times q}$分别是$m\times n$阶，$p\times q$阶常数矩阵，则

$$E(\boldsymbol{AXB})=\boldsymbol{A}(E\boldsymbol{X})\boldsymbol{B}$$

证明

$$\begin{aligned}E(\boldsymbol{AXB})&=E\left(\sum_{i=1}^{n}\sum_{j=1}^{p}a_{li}X_{ij}b_{jk}\right)\\&=\left(\sum_{i=1}^{n}\sum_{j=1}^{p}a_{li}E(X_{ij})b_{jk}\right)\\&=\boldsymbol{A}(E\boldsymbol{X})\boldsymbol{B}\end{aligned}$$

引理 2　若$\boldsymbol{A}$是 $m\times n$ 阶常数矩阵，$\boldsymbol{B}$ 是 $s\times r$ 阶常数矩阵，$\boldsymbol{X}'=(X_1, X_2, \cdots, X_n)$ 是 n 维随机向量，$\boldsymbol{Y}'=(Y_1, Y_2, \cdots, Y_r)$是 r 维随机向量，定义$\boldsymbol{X}$的方差矩阵为

$$D\boldsymbol{X}=E(\boldsymbol{X}-E\boldsymbol{X})(\boldsymbol{X}-E\boldsymbol{X})',$$

$\boldsymbol{X}$，$\boldsymbol{Y}$的协方差矩阵为

$$cov(\boldsymbol{X}, \boldsymbol{Y})=E(\boldsymbol{X}-E\boldsymbol{X})(\boldsymbol{Y}-E\boldsymbol{Y})'$$

则

$$D(AX)=A(DX)A'$$

$$cov((AX,\ BY)=Acov(X,\ Y)B'$$

证明

$$\begin{aligned}D(AX)&=E[(AX-E(AX))(AX-E(AX))']\\&=E[A(X-EX)(X-EX)'A']\\&=A[E(X-EX)(X-EX)']A'\\&=A(DX)A'\end{aligned}$$

$$\begin{aligned}cov(AX,\ BY)&=E[(AX-E(AX))(BY-E(BY))']\\&=E[A(X-EX)(Y-EY)'B']\\&=Acov(X,\ Y)B'\end{aligned}$$

定理5.1.1　若总体X服从正态分布$N(\mu,\ \sigma^2)$，$(X_1,\ X_2,\cdots,\ X_n)$是X的一个子样，则子样均值$\overline{X}$服从正态分布$N(\mu,\ \sigma^2/n)$。

证明　由第三章定理3.5.1和定理3.5.2知道，$\overline{X}$仍然服从正态分布，且

$$\begin{aligned}E(\overline{X})&=\frac{1}{n}E(X_1+X_2+\cdots+X_n)\\&=\frac{1}{n}[EX_1+EX_2+\cdots+EX_n]\\&=\frac{1}{n}n\mu\\&=\mu\end{aligned}$$

$$\begin{aligned}D(\overline{X})&=\frac{1}{n^2}D(X_1+X_2+\cdots+X_n)\\&=\frac{1}{n^2}[DX_1+DX_2+\cdots+DX_n]\\&=\frac{1}{n^2}n\sigma^2\\&=\sigma^2/n\end{aligned}$$

所以　$\overline{X}\sim N(\mu,\ \sigma^2/n)$。

定理5.1.2　设$X_1,\ \cdots,\ X_n$是取自正态总体$N(\mu,\ \sigma^2)$的一个

子样，其子样均值与子样方差分别为

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i \qquad S^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

则

1 $\overline{X}$与S^2相互独立，

2 $nS^2/\sigma^2\sim\chi^2(n-1)$。

证明 记 $X'=(X_1, X_2, \cdots, X_n)$，$Y'=(Y_1, Y_2, \cdots, Y_n)$ 且假定

$$Y=AX$$

其中

$$A=\begin{pmatrix}\frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}}\\ \frac{1}{\sqrt{2\cdot 1}} & -\frac{1}{\sqrt{2\cdot 1}} & 0 & \cdots & 0 & 0\\ \frac{1}{\sqrt{3\cdot 2}} & \frac{1}{\sqrt{3\cdot 2}} & \frac{-2}{\sqrt{3\cdot 2}} & \cdots & 0 & 0\\ \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots & \cdots\cdots\\ \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \cdots & \frac{1}{\sqrt{n(n-1)}} & \frac{-(n-1)}{\sqrt{n(n-1)}}\end{pmatrix}$$

是正交矩阵，即 $A'A=AA'=I_n$，且矩阵 A 从第二行起，每一行的元素之和为 0 。这时，由$Y=AX$ 知 $Y_1=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}X_i=\sqrt{n}\,\overline{X}$，从而$Y_1^2=n\overline{X}^2$。注意到

$$Y'Y=X'A'AX=X'I_nX=X'X=\sum_{i=1}^{n}X_i^2=\sum_{i=1}^{n}(X_i-\overline{X})^2$$

$$+n\overline{X}^2,\ \Rightarrow nS^2=\sum_{i=1}^{n}(X_i-\overline{X})^2=Y'Y-n\overline{X}^2$$

$$=\sum_{i=1}^{n}Y_i^2-Y_1^2=\sum_{i=2}^{n}Y_i^2$$

因为Y_j是诸独立正态随机变量X_i的线性和，$1\leqslant j\leqslant n$，从而诸Y_j也是正态变量，但

$$EY=E(AX)=AEX=A\cdot\mu 1=\mu A1,$$

这里$A1$是矩阵A与元素全是1的列向量1之积，即将A的各行元素分别求和，所以得

$$\mu A1=\begin{pmatrix}\sqrt{n}\,\mu\\ 0\\ \vdots\\ 0\end{pmatrix}$$

又

$$DY=D(AX)=A(DX)A'=A(\sigma^2 I_n)A'$$
$$=\sigma^2 AA'=\sigma^2 I_n$$

即诸Y_j，$j=1$，2，…，n的方差全都等于σ^2，且它们之间的协方差全为0。

在正态的情形下，不相关与相互独立是等价的，所以诸Y_j，$j=1$，2，…，n相互独立，从而Y_1与$\sum\limits_{j=2}^{n}Y_j^2$相互独立，于是$\overline{X}$与$S^2$相互独立。

现在，$\dfrac{nS^2}{\sigma^2}=\sum\limits_{j=2}^{n}\left(\dfrac{Y_j}{\sigma}\right)^2$而$\dfrac{Y_j}{\sigma}\sim N(0,1)$，$j=2$，…，$n$，所以$\dfrac{nS^2}{\sigma^2}$是$n-1$个正态$N(0,1)$变量的平方和．由第三章定理3.5.9知，它服从自由度为$n-1$的χ^2分布。

推论1　设X_1，X_2，…，X_n是取自正态总体$N(\mu,\ \sigma^2)$的一个子样，$\overline{X}$与S^2分别为它的子样均值与子样方差，则

$$T=\frac{\overline{X}-\mu}{\dfrac{S}{\sqrt{n-1}}}$$

服从自由度为$n-1$的t分布。

证明　由定理 5.1.1 知，$\overline{X} \sim N(\mu, \sigma^2/n)$，所以，$\overline{X}$的标准化

$$\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$$

但 $nS^2/\sigma^2 \sim \chi^2(n-1)$，注意到

$$T=\frac{\overline{X}-\mu}{\frac{S}{\sqrt{n-1}}}=\frac{\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{nS^2}{(n-1)\sigma^2}}}=\frac{\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}}{\sqrt{\frac{nS^2}{\sigma^2}}}\sqrt{n-1}$$

由第三章定理3.5.10知它服从自由度为 $n-1$ 的 t 分布。

推论 2　设 $X_1, X_2, \cdots, X_{n_1}$ 和 $Y_1, X_2, \cdots, Y_{n_2}$ 分别是从正态总体$N(\mu_1, \sigma_1^2)$与 $N(\mu_2, \sigma_2^2)$抽取的两个相互独立的子样，并设

$$S_1^2=\frac{1}{n_1}\sum_{i=1}^{n_1}(X_i-\overline{X})^2$$

$$S_2^2=\frac{1}{n_2}\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2$$

分别为这两个子样的子样方差，其中$\overline{X}=\frac{1}{n_1}\sum_{i=1}^{n_1}X_i$，$\overline{Y}=\frac{1}{n_2}\sum_{i=1}^{n_2}Y_i$，则

$$F=\frac{n_1S_1^2}{n_2S_2^2}\cdot\frac{(n_2-1)\sigma_2^2}{(n_1-1)\sigma_1^2}$$

服从$F(n_1-1, n_2-1)$分布。

特别地，当 $\sigma_1^2=\sigma_2^2$ 时，则

$$F=\frac{n_1(n_2-1)S_1^2}{n_2(n_1-1)S_2^2}$$

服从 $F(n_1-1, n_2-1)$ 分布。

这个结论由定理5.1.2与第三章定理3.5.11直接推出。

推论 3　设 $X_1, X_2, \cdots, X_n$ 是取自正态总体 $N(\mu_1, \sigma^2)$的子

样，Y_1，Y_2，…，Y_{n_2}是取自正态总体$N(\mu_2, \sigma^2)$的子样，并且它们相互独立，则随机变量

$$\frac{(\overline{X}-\overline{Y})-(\mu_1-\mu_2)}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

服从自由度为 n_1+n_2-2 的 t 分布，其中$\overline{X}=\frac{1}{n_1}\sum_{i=1}^{n_1}X_i$，$\overline{Y}=\frac{1}{n_2}\sum_{i=1}^{n_2}Y_i$，$S_1^2=\frac{1}{n_1}\sum_{i=1}^{n_1}(X_i-\overline{X})^2$，$S_2^2=\frac{1}{n_2}\sum_{i=1}^{n_2}(Y_i-\overline{Y})^2$，$S^2=\frac{1}{n_1+n_2-2}(n_1S_1^2+n_2S_2^2)$。

证明　注意到

$$E(\overline{X}-\overline{Y})=E\overline{X}-E\overline{Y}=\mu_1-\mu_2$$

$$D(\overline{X}-\overline{Y})=D\overline{X}+D(\overline{Y})=\left(\frac{1}{n_1}+\frac{1}{n_2}\right)\sigma^2$$

因此

$$\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}\sim N(0,\ 1)$$

由定理 5.1.2 知

$$\frac{n_1S_1^2}{\sigma^2}\sim\chi^2(n_1-1),\qquad \frac{n_2S_2^2}{\sigma_2}\sim\chi^2(n_2-1),$$

且它们相互独立，由第三章定理3.5.8关于独立的χ^2变量的可加性知

$$\frac{n_1S_1^2}{\sigma^2}+\frac{n_2S_2^2}{\sigma^2}$$

是 $\chi^2(n_1+n_2-2)$ 变量，从而得出

$$\frac{\dfrac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sigma\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}}{\sqrt{\dfrac{n_1S_1^2+n_2S_2^2}{\sigma^2(n_1+n_2-2)}}}=\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}}$$

服从自由度为 n_1+n_2-2 的 t 分布。

§5.2 点 估 计

参数估计是统计推断的一个基本问题。在统计推断理论中，对均值、方差等未知参数进行估计叫做参数估计；对概率分布进行估计叫做非参数估计。

对参数进行估计有两种方法：一是点估计，一是区间估计。这一节先讲点估计。

什么叫点估计？我们仍用灯泡寿命的例子来说明。为了了解灯泡的质量，自然应当对灯泡寿命X的均值 EX 及方差 DX 进行估计。为此，从一大批灯泡中随机抽取 n 个，得到它们的寿命值为 X_1，X_2，…，X_n。由大数定律知道，当 n 很大时，$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 以很大的概率接近于 EX。所以，我们自然会想到用子样均值$\overline{X}$作为 EX 的估计。由于 DX 实际上是 $(X-EX)^2$ 的均值，所以用子样方差$S^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 去估计 DX 也是十分自然的。

一般地说，如果 θ 是X的一个未知参数，用子样 $(X_1, X_2, \cdots, X_n)$ 的某个函数$\hat{\theta}(X_1, X_2, \cdots, X_n)$ 去估计 θ，就称为对 θ 作点估计。这里，用来估计总体的未知参数 θ 的子样函数 $\hat{\theta}(X_1, X_2, \cdots, X_n)$ 称为参数 θ 的估计量。以后，凡是估计量都用在该参数上标记符号“^”表示。估计量是随机变量，其中不应含有任何未知参数。如果 $(x_1, x_2, \cdots, x_n)$是子样的一组观测值，代入估计量$\hat{\theta}$后就得到$\hat{\theta}$的一个具体的数值，通常称这一数值为 θ 的估计值。今后，我们将不再强调估计量与估计值这两个名词的区别，统称为估计。

由于估计量是随机变量，因而给出一组子样观测值 $(x_1, x_2, \cdots, x_n)$，相应地就得到 θ 的一个估计值。对于不同的子样观测值，

所得到的估计值则是不同的。例如，对灯泡寿命 EX 进行点估计，相应于子样的一组观测值，我们得到子样均值的一个值，用它作为 EX 的估计值。另作一批抽样，又得到 $(X_1,\cdots,X_n)$ 的另一组观测值，它的算术平均值是子样均值这一随机变量所取的另一数值。它也是 EX 的一个估计值。由此，我们可以看出，寻找一个估计量实质上是找依据观测结果去求得未知参数 θ 的估计值的一种方法，而不是具体地去求出一个估计值。衡量一个估计的优劣当然不能由一次试验的结果作出定论，而必须从全局上，从多次由试验结果得到的估计值与参数真值的偏差大小来确定。这就是一个估计量的优良性问题。稍后，我们就将详加讨论。

（一）求估计量的方法

在这一段里，我们讲两种常用的求估计量的方法，一是矩法，一是最大似然法。

1　矩法

矩法就是用子样矩估计总体矩，从而得到未知参数估计量的方法。比如，设 X 的分布为 $p(x;\ \theta_1,\ \cdots,\ \theta_k)$，假定 X 的 k 阶矩 m_k 存在，则 m_r 是 $\theta_1,\ \cdots,\ \theta_k$ 的函数 $(r=1,\ 2,\ \cdots,\ k)$。根据子样 $X_1,\ X_2,\ \cdots,X_n$，求子样 r 阶矩 $\widehat{m}_r$，令

$$m_r(\theta_1,\ \cdots,\ \theta_k)=\widehat{m}_r\quad r=1,\ 2,\ \cdots,\ k \tag{5.2.1}$$

则从这 k 个方程中，可解得 $\widehat{\theta}_r(X_1,\ \cdots,\ X_n)$，$r=1,\ 2,\ \cdots,\ k$。这样得到的估计量叫做矩法估计。

例5.2.1　总体均值与方差的矩法估计

解　设子样为 $(X_1,\ X_2,\ \cdots,\ X_n)$ 且总体 X 的二阶矩存在，因此有 $m_2=\sigma^2+m_1^2$。用矩法得方程组

$$\begin{cases}\widehat{m}_1=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}X_i=\overline{X},\\ \widehat{\sigma}^2+\widehat{m}_1^2=\widehat{m}_2=\dfrac{1}{n}\displaystyle\sum_{i=1}^{n}X_i^2\end{cases}$$

解之得

$$\hat{m}_1 = \overline{X}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2 = S^2$$

所以，子样均值$\overline{X}$与子样方差S^2分别是总体均值m_1与总体方差σ^2的矩法估计。

例5.2.2　设X服从均匀分布

$$p(x,\theta) = \begin{cases} \frac{1}{\theta_2} & x \in [\theta_1, \theta_1 + \theta_2] \\ 0 & x \notin [\theta_1, \theta_1 + \theta_2] \end{cases}$$

求θ_1，θ_2的矩法估计量

解　$m_1 = EX = \theta_1 + \frac{\theta_2}{2}$

$$m_2 = EX^2 = \int_{\theta_1}^{\theta_1+\theta_2} \frac{1}{\theta_2} x^2 dx = \theta_1^2 + \theta_1\theta_2 + \frac{\theta_2^2}{3}$$

由（5.2.1）得两个方程

$$\theta_1 + \frac{\theta_2}{2} = \overline{X}$$

$$\theta_1^2 + \theta_1\theta_2 + \frac{\theta_2^2}{3} = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \qquad (5.2.2)$$

解（5.2.2）得

$$\hat{\theta}_2 = \sqrt{12}\left\{\frac{1}{n}\sum_{i=1}^{n} X_i^2 - (\overline{X})^2\right\}^{\frac{1}{2}} = \sqrt{12}S$$

$$\hat{\theta}_1 = \overline{X} - \frac{\hat{\theta}_2}{2}$$

2　最大似然法

设总体X服从分布$p(x;\theta_1,\cdots,\theta_k)$（当$X$为连续型时为概率密度，当$X$为离散型时为概率函数），其中$\theta_1,\cdots,\theta_k$为$k$个未知参数。若$x_1,x_2,\cdots,x_n$为子样$X_1,X_2,\cdots,X_n$的观测值，则

$$L(x_1, x_2, \cdots, x_n; \theta_1, \cdots, \theta_k) = \prod_{i=1}^{n} p(x_i; \theta_1, \cdots, \theta_k)$$

(5.2.3)

是子样 $X_1, X_2, \cdots, X_n$ 取值为 $x_1, x_2, \cdots, x_n$的概率（连续型时为概率密度），它是 $\theta_1, \cdots, \theta_k$ 的函数，称为 $\theta_1, \cdots, \theta_k$ 的似然函数，简记为 $L(\theta)$（θ 代表 $\theta_1, \cdots, \theta_k$）。最大似然法就是把似然函数达到最大值的$\hat{\theta}_1, \hat{\theta}_2, \cdots, \hat{\theta}_k$作为$\theta_1, \cdots, \theta_k$的估计值。称 $\hat{\theta}_i$为θ_i的最大似然估计量。

由于对数函数的单调性，所以使$L(\theta)$达到最大的$\hat{\theta}_i$与使 $\ln L(\theta)$ 达到最大值的$\hat{\theta}_i$是相同的。这样只要求出使 $\ln L(\theta)$达到最大的 $\hat{\theta}_i$ 就行了。因此，可从

$$\frac{\partial \ln L(\theta)}{\partial \theta_i} = 0, \quad i = 1, \cdots, k \qquad (5.2.4)$$

求 $\hat{\theta}_i$。

最大似然法的道理是显然的，因为在子样的一次观测中就出现了 $x_1, x_2, \cdots, x_n$，这说明 $\theta_1, \cdots, \theta_k$ 的值很可能是那样的值，以它为分布的参数时能使得 $x_1, x_2, \cdots, x_n$ 最容易出现。因此，我们估计 $\theta_1, \cdots, \theta_k$ 的值为使 $x_1, \cdots, x_n$ 出现概率（或密度）$L(\theta)$ 达到最大的值。

例 5.2.3 已知总体X服从均匀分布

$$p(x;\theta) = \begin{cases} \dfrac{1}{\theta} & 0 < x \leqslant \theta \\ 0 & 其它 \end{cases}$$

求 θ 的最大似然估计量。

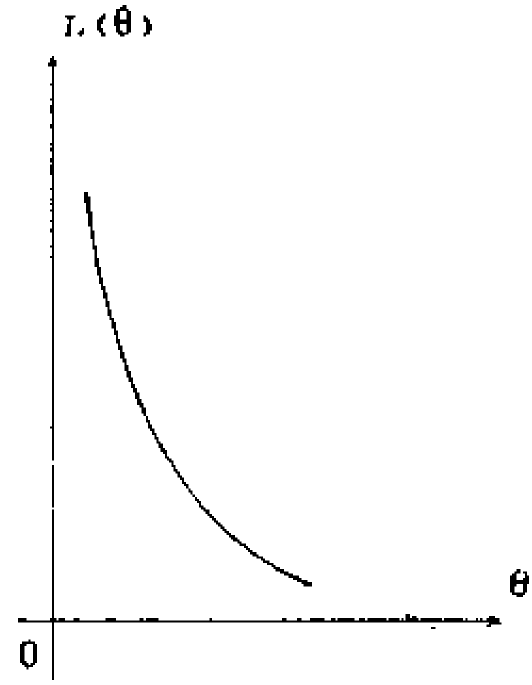
解 似然函数为

$$L(\theta) = \prod_{i=1}^{n} p(x_i;\theta) = \begin{cases} 1/\theta^n & (0 < x_i \leqslant \theta) \quad i = 1, 2, \cdots, n \\ 0 & 其他 \end{cases}$$

$L(\theta)$ 的图形如右图5.2.1，在 $\theta>0$ 时，单调减小，θ 愈小，则 $L(\theta)$ 愈大，但 θ 又必须不小于所有的 x_i，$\theta \geqslant x_i$　$i=1, 2, \cdots, n$，所以 $\theta = \max\limits_{1\leqslant i\leqslant n} x_i$ 时，$L(\theta)$时达到最大值，因此

$$\hat{\theta} = \max_{1\leqslant i\leqslant n} X_i$$

是 θ 的最大似然估计量。



图（5.2.1）

例 5.2.4　已知总体X服从正态分布

$$p(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \qquad (-\infty<x<\infty)$$

求 μ，σ^2 的最大似然估计。

解　似然函数为

$$L(\theta) = L(\mu, \sigma^2) = \frac{1}{(\sigma\sqrt{2\pi})^n} e^{-\sum\limits_{i=1}^{n} \frac{(x_i-\mu)^2}{2\sigma^2}}$$

$$\ln L(\mu, \sigma^2) = -\frac{n}{2}\{\ln(2\pi) + \ln(\sigma^2)\} - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(X_i-\mu)^2$$

$$\frac{\partial \ln L(\mu, \sigma^2)}{\partial\mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(X_i-\mu) = \frac{1}{\sigma^2}\left\{\sum_{i=1}^{n} X_i - n\mu\right\} = 0$$

$$\frac{\partial \ln L(\mu, \sigma^2)}{\partial\sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(X_i-\mu)^2 = 0$$

解上面方程得 $\hat{\mu} = \overline{X}$，$\hat{\sigma}_2 = \frac{1}{n}\sum\limits_{i=1}^{n}(X_i-\mu)^2$ 由于

$$\frac{\partial^2 \ln L(\mu, \sigma^2)}{\partial\mu^2} = -\frac{n}{\sigma^2} < 0$$

所以$\overline{X}$，S^2 分别为μ和σ^2的最大似然估计量。

（二）估计量的优良性准则

上面讲到可以用子样均值$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 去估计 EX（简写为μ），除此之外，也还可以用其他的估计量去估计μ，比如也可以用$\hat{\mu}_1=\sum_{i=1}^{n}\alpha_iX_i$去估计$\mu$，其中$\alpha_i$满足$\alpha_i\geqslant 0$，$i=1,2,\cdots,n$，且$\sum_{i=1}^{n}\alpha_i=1$。这样就有一个问题，即哪个估计量好？我们总是希望用一个最好的估计量去作估计。这里要注意，所谓估计的好坏、优劣不是指估计量的一次观测值来说的，因为，比如在某个子样里，可能$\overline{X}$比$\hat{\mu}_1$接近于μ，而在另一个子样里，$\hat{\mu}_1$可能比$\overline{X}$更接近于μ，所以估计的好与坏是对估计量来说的。

什么样的估计量是好的估计量呢？我们这里举出几条衡量估计量是否优良的准则。

（1）无偏性

定义 5.2.1 若未知参数θ的估计量$\hat{\theta}(X_1, X_2, \cdots, X_n)$满足下式

$$E(\hat{\theta})=\theta$$

则称$\hat{\theta}$具有无偏性，称$\hat{\theta}$为θ的无偏估计量。

例 5.2.5 子样均值

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$$

是总体均值μ的无偏估计量。

证明 由于$E(X_i)=\mu$，$i=1, 2, \cdots, n$

所以 $$E(\overline{X})=\frac{1}{n}E\left(\sum_{i=1}^{n}X_i\right)=\frac{1}{n}\sum_{i=1}^{n}EX_i=\mu$$

例5.2.6 子样方差

$$S^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

不是总体方差DX（简写为σ^2）的无偏估计量．而

$$S^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{n}{n-1}S^2$$

是σ^2的无偏估计量．

证明

$$\begin{aligned}
E(S^2)&=E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2\right]\\
&=\frac{1}{n}E\left[\sum_{i=1}^{n}(X_i-\mu+\mu-\overline{X})^2\right]\\
&=\frac{1}{n}E\left[\sum_{i=1}^{n}(X_i-\mu)^2-2\sum_{i=1}^{n}(X_i-\mu)(\overline{X}-\mu)\right.\\
&\qquad\left.+n(\overline{X}-\mu)^2\right]\\
&=\frac{1}{n}E\left[\sum_{i=1}^{n}(X_i-\mu)^2-n(\overline{X}-\mu)^2\right]\\
&=\frac{1}{n}\sum_{i=1}^{n}E(X_i-\mu)^2-E(\overline{X}-\mu)^2\\
&=\frac{1}{n}\sum_{i=1}^{n}DX_i-D\overline{X}\\
&=\frac{1}{n}n\sigma^2-\frac{\sigma^2}{n}\\
&=\frac{n-1}{n}\sigma^2
\end{aligned}$$

$$E(S^{*2})=E\left(\frac{n}{n-1}S^2\right)=\frac{n}{n-1}\cdot\frac{n-1}{n}\sigma^2=\sigma^2$$

由此，顺便得出

$$E(nS^2)=E\left\{\sum_{i=1}^{n}(X_i-\overline{X})^2\right\}=(n-1)\sigma^2 \qquad (5.2.5)$$

请读者注意，这一等式以后将多次用到．

无偏性说明当$\hat{\theta}$是θ的无偏估计量时，虽然$\hat{\theta}$的每次观测值可能与θ不同，但是多次用$\hat{\theta}$去估计θ时，估计值的平均接近于θ。这当然是一个较好的性质．但是，只根据无偏性来确定估计量还是不够的，因为虽然 $E(\hat{\theta})=\theta$，但若 $D(\hat{\theta})$很大，即$\hat{\theta}$对θ 的偏离程度很大，这个估计量也还不是好的．因此，又有有效性的要求．

（2）有效性

定义5.2.2 若$\hat{\theta}_1$和$\hat{\theta}_2$都是θ的无偏估计量，而且$D(\hat{\theta}_1)\leqslant D(\hat{\theta}_2)$，则称$\hat{\theta}_1$比$\hat{\theta}_2$有效．

例 5.2.7 μ的估计量$\hat{\mu}_1=\overline{X}$比$\hat{\mu}_2=\sum\limits_{i=1}^{n}\alpha_iX_i\left(\alpha_i\geqslant 0,\ i=1,2,\cdots,n,\ \sum\limits_{i=1}^{n}\alpha_i=1\right)$有效．

证明　由于

$$E(\hat{\mu}_2)=E\left(\sum_{i=1}^{n}\alpha_iX_i\right)=\sum_{i=1}^{n}\alpha_iEX_i=\mu$$

所以$\hat{\mu}_2$也是μ的无偏估计量，但是由于许瓦兹(Schwarz)不等式

$$\left(\sum_{i=1}^{n}x_iy_i\right)^2\leqslant\left(\sum_{i=1}^{n}x_i^2\right)\left(\sum_{i=1}^{n}y_i^2\right)$$

知

$$\left(\sum_{i=1}^{n}\alpha_i\right)^2\leqslant n\sum_{i=1}^{n}\alpha_i^2$$

从而

$$D(\hat{\mu}_1)=\frac{\sigma^2}{n}=\frac{\sigma^2}{n}\left(\sum_{i=1}^{n}\alpha_i\right)^2\leqslant\sigma^2\left(\sum_{i=1}^{n}\alpha_i^2\right)=D(\hat{\mu}_2)$$

所以$\overline{X}$比$\sum\limits_{i=1}^{n}\alpha_iX_i$有效．

例5.2.8 若总体X服从普阿松分布

$$P(X=k)=\frac{e^{-\lambda}}{k!}\lambda^k\qquad k=0,1,2\cdots$$

对于容量为$n\ (>2)$的子样$X_1, X_2, \cdots, X_n$，$\hat{\lambda}_1 = \overline{X}$比$\hat{\lambda}_2 = (X_1 + X_2)/2$有效。

证明　由于$E(X_i) = \lambda$，所以$E(\hat{\lambda}_1) = \lambda$，$E(\hat{\lambda}_2) = \lambda$，即$\hat{\lambda}_1$与$\hat{\lambda}_2$都是$\lambda$的无偏估计量。但是$D(\hat{\lambda}_1) = \dfrac{\lambda}{n}$，而$D(\hat{\lambda}_2) = \dfrac{\lambda}{2}$，所以$\hat{\lambda}_1$比$\hat{\lambda}_2$有效。

定义 5.2.3　对于固定的子样容量n，在参数θ的一切无偏估计量中，如果$D(\hat{\theta})$的值达到极小，则称$\hat{\theta}$为θ的最小方差无偏估计量。

判断一个估计量是否是最小方差无偏估计量，这个问题比较复杂，不准备详细介绍，这里只给出以下结果。

（a）频率$\dfrac{k}{n}$是概率p的最小方差无偏估计量。（b）若总体X服从正态分布$N(\mu, \sigma^2)$，则$\overline{X}$与S^{*2}分别是μ与σ^2的最小方差无偏估计量。

（3）一致性

定义5.2.4　设$\hat{\theta}(X_1, X_2, \cdots, X_n)$是$\theta$的估计量，如果对于任意$\varepsilon > 0$，当$n \to \infty$时，有

$$P(|\hat{\theta} - \theta| \geqslant \varepsilon) \to 0 \tag{5.2.6}$$

则称$\hat{\theta}$为θ的一致估计量。若（5.2.6）成立，则说$\hat{\theta}$依概率收敛于θ，因此也可说如果当$n \to \infty$时$\hat{\theta}$依概率收敛于θ，则$\hat{\theta}$称为θ的一致估计量。

例 5.2.9　频率$\dfrac{k}{n}$是概率p的一致估计量。

因为由大数定理已知，对任意的$\varepsilon > 0$有

$$P\left(\left|\frac{k}{n} - p\right| \geqslant \varepsilon\right) \to 0 \qquad (n \to \infty)$$

例 5.2.10　子样均值$\overline{X}$是总体均值μ的一致估计，而$S^{*2} =$

$\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 及 $S^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 都是总体方差 σ^2 的一致估计量．

证明　根据大数定理已知，对任意 $\varepsilon>0$ 有

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n}X_i-\mu\right|\geqslant\varepsilon\right)\to 0\qquad(n\to\infty)$$

所以 $\overline{X}$ 是 μ 的一致估计量．

其次由于

$$S^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2=\frac{1}{n}\sum_{i=1}^{n}(X_i^2-2X_i\overline{X}+\overline{X}^2)$$

$$=\frac{1}{n}\sum_{i=1}^{n}X_i^2-2(\overline{X})^2+(\overline{X})^2=\overline{X^2}-(\overline{X})^2$$

由大数定律可知

$$\overline{X^2}=\frac{1}{n}\sum_{i=1}^{n}X_i^2\xrightarrow{P}E(X^2)\quad(\xrightarrow{P}\text{表示依概率收敛})$$

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i\xrightarrow{P}EX$$

根据依概率收敛的性质，则有

$$S^2=\overline{X^2}-(\overline{X})^2\xrightarrow{P}E(X^2)-(EX)^2=\sigma^2$$

所以 S^2 是 σ^2 的一致估计量．

又由于 $\frac{n}{n-1}\longrightarrow 1$　所以 $S^{*2}=\frac{n}{n-1}S^2$ 也是 σ^2 的一致估计量．

§5.3　显著性检验

什么叫显著性检验？它包括一些什么样的内容？为了弄清这些问题，我们先从什么是假设检验谈起．

举两个例子来说明

例 5.3.1 某车间所生产的铆钉，根据以往的生产情况，已知其平均直径$\mu_0=2$，标准差 $\sigma_0=0.1$（厘米）。现在又采用一种新工艺，在用新工艺生产的铆钉中抽取了$n=100$个，测得其平均直径为$\overline{x}=1.978$（厘米）。我们知道，既使工艺条件不变，所生产的铆钉的直径也不会全等于 $\mu_0=2$，而是在μ_0附近波动。因此自然要问，$\overline{x}$与μ_0的差异是纯属偶然波动，还是反映了工艺改变的影响？也就是要问，采用新工艺前后所生产的铆钉的平均直径是否相同。如用μ表示采用新工艺后的铆钉的平均直径，即要求判断 $\mu=\mu_0$ 是否成立。

例 5.3.2 某种汽车零件，按新工艺投产后，想了解其生产情况是否正常，为此需要了解其某一指标X是否服从正态分布。于是在大批产品中抽取了n件，如何根据X的n个观测值 x_1，$x_2\cdots$，x_n判断X是否服从正态分布。若用 $F(x)$ 表示X的概率分布，即要求判断 $F(x)=\Phi(x)$是否成立。

在这两个问题里，都有一个关于总体的假设，一般用 H_0 表示。在例5.3.1中 H_0 是 $\mu=\mu_0$，在例 5.3.2 中 H_0是$F(x)=\Phi(x)$，称它们为统计假设。所考虑的问题就是如何根据子样，对假设 H_0 作出判断，判断它成立还是不成立。这就是假设检验问题。

如果统计假设是关于总体参数的，如例 5.3.1，就是参数检验问题。如果统计假设是关于总体分布的，如例 5.3.2，就是分布检验问题。

在上述这些问题中，只提出了一个统计假设。但是，在实际问题中，往往还需要提出两个甚至多个统计假设，而且检验的目的也需要同时判断多个假设中哪一个成立。如果在一个统计问题中仅提出一个统计假设，而且我们的目的也仅仅是判断这一个统计假设是否成立，并不同时研究其它统计假设。这类检验问题称为显著性检验。本节主要讨论正态总体参数的显著性检验问题。

怎样去检验一个统计假设呢？在展开一般的讨论之前，我们讲一下它的基本想法。

假设检验方法的主要依据是实际推断原理，用一句话来说，就

是："小概率事件在一次试验中实际上是不可能发生的"。

举个例子来说明，假如在1000粒大豆中含有一粒虫豆，现在从中随机取一粒，则"取得虫豆"的概率很小，是1/1000，因此，在1000粒大豆中只含有一粒虫豆的假设下，从中只取一粒，就正好取得虫豆，实际上（不是理论上）是不可能的。

根据这个原理，我们就得到一个推理的方法，即，如果在某假设成立的条件下，事件A是一个小概率事件，现在只进行一次试验，如果在一次试验中A就发生了，则自然有理由认为原来的假设是不成立的。比如，如果在1000粒大豆中，随机取一粒，结果就是虫豆，则自然可以认为1000粒大豆中的虫豆不止一粒，即虫豆率 p 不是1/1000。这就是进行假设检验的基本思想方法。

现在结合例5.3.1说明在假设检验中怎样用这个推理想法。

在例 5.3.1 中，我们看到$\overline{x}$与μ_0存在差异。常识告诉我们，若$\mu=\mu_0$成立，则这个差异$|\overline{x}-\mu_0|$比较大的可能性小，而$|\overline{x}-\mu_0|$比较小的可能性大。因此，我们可以根据$|\overline{x}-\mu_0|$的大小来判断H_0是否成立，这样就需要分析在H_0成立时，$|\overline{x}-\mu_0|$大到什么程度，或说大于一个什么值，是一个小概率事件。这当然要根据$|\overline{X}-\mu_0|$的分布来求。如果这个值找到了，比如已知$P(|\overline{X}-\mu_0|>c)=\alpha$（$\alpha$为很小的值，一般取为0.05，0.01等），那么我们就可以根据观测值计算$|\overline{x}-\mu_0|$，若$|\overline{x}-\mu_0|>c$，即小概率事件发生了，就可以认为H_0不成立，若$|\overline{x}-\mu_0|<c$，就不能否定H_0。下一段讲的各种参数检验问题，虽然具体方法各有不同，但都是根据这个想法。

（一）关于正态总体均值 μ 的显著性检验

1　在 $\sigma=\sigma_0$ 已知时，检验 H_0：$\mu=\mu_0$

由上节分析可知关键在于求出当H_0成立时$\overline{X}-\mu_0$ 的分布。由于当H_0成立时X服从分布$N(\mu_0,\sigma_0^2)$，所以$\overline{X}$服从分布$N\left(\mu_0,\ \frac{\sigma_0^2}{n}\right)$，

因而$\overline{X}-\mu_0$服从分布$N\left(0,\frac{\sigma_0^2}{n}\right)$。为了求值方便，将其标准化，考虑量$U=\frac{\overline{X}-\mu_0}{\sigma_0}\sqrt{n}$，它服从分布 $N(0,1)$。

对于已给的小概率α，找出满足下式的u_α

$$P(|U|\geqslant u_\alpha)=\alpha$$

如果U的观测值$|u|\geqslant u_\alpha$，则拒绝假设H_0，如果$|u|<u_\alpha$，则可考虑接受H_0。

这种检验法称为U检验法。

在例 5.3.1 中，假定已给$\alpha=0.05$，则$u_\alpha=1.96$，$\overline{x}=1.978$，$\mu_0=2$，所以

$$|u|=\left|\frac{\overline{x}-\mu_0}{\sigma}\sqrt{n}\right|=\left|\frac{1.978-2}{0.1}\sqrt{100}\right|=2.2>1.96$$

所以拒绝假设H_0。即认为采用新工艺后，平均直径不再等于 2 了。

回顾一下（1）的检验方法，我们看到量U的选择及其分布的求出是关键。量U是子样的函数，并且把在子样中所包含的我们关心的信息集中起来，这样的函数在数理统计学中称为统计量。严格地说，一个统计量就是子样$(X_1,X_2,\cdots,X_n)$的一个（可测）函数，并且要求它不包含有任何未知参数。因此，它也是一个随机变量。由统计量的定义可知，估计量也是一种统计量。

概括一下，显著性检验问题的处理步骤是：

（1）建立统计假设H_0；

（2）构造一个合适的统计量U并由子样观测值算出统计量U的观测值u；

（3）规定一个小概率α（一般取$\alpha=0.05$或$\alpha=0.01$），求出在H_0成立的条件下，满足

$$P(|U|\geqslant u_\alpha)=\alpha$$

的临界值u_α，以后，α称为显著水平；

（4）比较观测值u和u_α，如果$|u|\geqslant u_\alpha$，则拒绝假设H_0。

区域$(-\infty, -u_\alpha)\cup(u_\alpha, +\infty)$称为拒绝域。

上述程序也适用于其它一些检验问题。通常，我们按检验时所取子样容量的大小，分为小子样与大子样两类问题。对于小子样的显著性检验，需要给出检验统计量U的精确分布，对于大子样问题可利用U的极限分布作为近似。

2　在σ未知时，检验H_0：$\mu=\mu_0$

当σ未知时，U检验法不能用，因为在σ未知时，算不出$U=\dfrac{\overline{X}-\mu_0}{\sigma}\sqrt{n}$的观测值来，为了解决这一矛盾，自然容易想到用子样方差

$$S^{*2}=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2$$

去代替未知的方差σ^2。这时，由定理5.1.2的推论1知，在假设H_0：$\mu=\mu_0$成立时，统计量

$$T=\frac{\overline{X}-\mu}{\dfrac{S^*}{\sqrt{n}}}=\frac{\overline{X}-\mu}{\sqrt{\dfrac{nS^2}{n(n-1)}}}=\frac{\overline{X}-\mu}{\dfrac{S}{\sqrt{n-1}}}$$

服从自由度为$n-1$的t分布。对于已给的α，找出满足下式的t_α

$$P(|T|\geqslant t_\alpha)=\alpha$$

根据子样$x_1, \cdots, x_n$，计算T的观测值t，若$|t|\geqslant t_\alpha$，则拒绝假设H_0，若$|t|<t_\alpha$，则不能拒绝H_0。

这种检验法叫t检验法。

当子样容量为n时，T统计量的自由度是$n-1$，这一点在查表时要注意，表中的n是自由度。

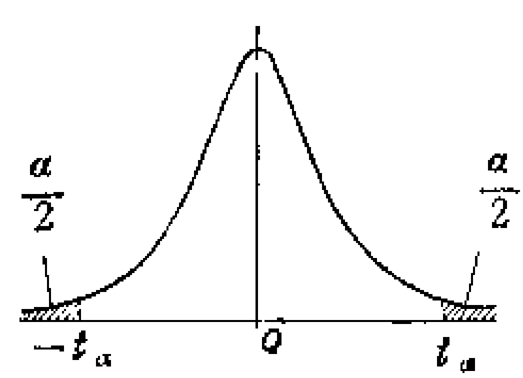


图 (5.3.1)

例5.3.3　已知甲地区每年的最高冻深平均为190厘米。今有乙地区从1954到1963年的每年最高冻深资料为156，156，185，185，205，209，188，190，205，181。问乙地区的平均最高冻深是否与

甲地区相同。($\alpha=0.05$)。

解　可认为乙地区的平均最高冻深X服从正态分布。这里要检验的统计假设是 H_0: $\mu=190$，由于σ未知，所以用t检验法。

这里 $n=10$，自由度为9，对 $\alpha=0.05$，查t分布表得 $t_\alpha=2.262$，即

$$P(|T|\geqslant 2.262)=0.05$$

根据资料计算t的观测值，为此先计算$\overline{x}$ 与 s^{*2}，$\overline{x}=186$，$\overline{s}^{*2}=\frac{1}{9}\sum_{i=1}^{10}(x_i-\overline{x})^2=342$，$s^*=18.6$，最后算得 $t=-0.683$，因为 $|-0.683|<2.263$，所以不能拒绝假设$\mu=190$，可以认为乙地区的平均最高冻深与甲地区的平均最高冻深无显著差异。

将具有自由度30的t分布表与正态分布表比较，可以看到它们几乎一致。这是因为当自由度数目无限增大时，t分布收敛于正态分布，即下面定理成立。

定理5.3.1　具有自由度n的t分布的分布函数序列$\{F_n(t)\}$满足关系式

$$\lim_{n\to\infty}F_n(t)=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{t}e^{-t^2/2}dt$$

证明从略。

现在考虑比较两个正态总体$N(\mu_1,\sigma_1^2)$,$N(\mu_2,\sigma_2^2)$均值的问题。当σ_1^2，σ_2^2为已知时　可用U检验法检验假设H_0: $\mu_1=\mu_2$。用统计量

$$U=\frac{\overline{X}-\overline{Y}}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$$

它服从正态分布 $N(0,1)$。

当σ_1^2，σ_2^2未知时，如果已知$\sigma_1=\sigma_2$（具体数值未知），用t检验法检验H_0: $\mu_1=\mu_2$。用统计量

$$T=\frac{\overline{X}-\overline{Y}}{\sqrt{\frac{(n_1S_1^2+n_2S_2^2)(n_1+n_2)}{n_1n_2(n_1+n_2-2)}}}$$

已知当 $\mu_1=\mu_2$ 成立时它服从自由度为 (n_1+n_2-2) 的 t 分布。

例 5.3.4　为了比较两个春小麦品种新曙光 1 号与丰强 2 号的产量，各在四块试验田上进行对比试验，结果产量如下（斤/亩）：

新曙光 4 号 x：475，449，407，441。

丰强 2 号 y：395，349，379，327。

以显著水平 $\alpha=0.01$ 检验两个品种的平均亩产是否相等。

解　可以认为两个品种的亩产量分别服从正态分布 $N(\mu_1,\sigma^2)$，$N(\mu_2,\sigma^2)$，方差相等。现在用 t 检验法检验 H_0：$\mu_1=\mu_2$。

这里 $m=n=4$，　　$\overline{x}=441$，　　$\overline{y}=362.5$

$$3S_1^2=\sum_{i=1}^{4}(x_i-\overline{x})^2=2313,\quad 3S_2^2=\sum_{i=1}^{4}(y_i-\overline{y})^2=2771$$

计算 T 的观测值得

$$|t|=78.5\sqrt{\frac{6\times16}{8\times5084}}\doteq3.81$$

对于 $\alpha=0.01$，自由度 6，查 t 分布表得 $t_\alpha=3.707$，由于 $|t|>t_\alpha$，所以可认为两个品种的平均亩产有显著差异。

（二）关于正态总体方差 σ^2 的显著性检验

（1）在 $\mu=\mu_0$ 已知时，检验 H_0：$\sigma^2=\sigma_0^2$。

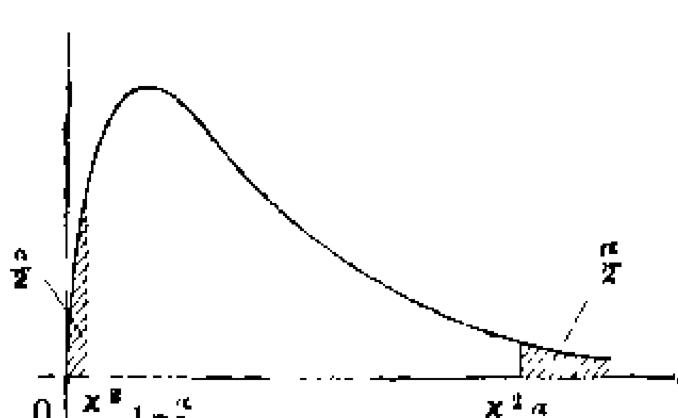
用统计量

$$\chi^2=\sum_{i=1}^{n}\left(\frac{X_i-\mu_0}{\sigma}\right)^2$$

当 $\sigma=\sigma_0$ 时，它服从自由度为 n 的 χ^2 分布。求 χ_1^2，χ_2^2 分别满足

$$P(\chi^2\geqslant\chi_2^2)=\frac{\alpha}{2}$$

$$P(\chi^2\geqslant\chi_1^2)=1-\frac{\alpha}{2}$$



图（5.3.2）

一般χ_1^2写为$\chi^2_{1-\frac{\alpha}{2}}$，$\chi^2_2$写为$\chi^2_{\frac{\alpha}{2}}$，则拒绝域为

$$(0,\ \chi^2_{1-\frac{\alpha}{2}})\text{和}(\chi^2_{\frac{\alpha}{2}},\ \infty).$$

若χ^2的观测值$\chi^2 \leqslant \chi^2_{1-\frac{\alpha}{2}}$ 或$\chi^2 \geqslant \chi^2_{\frac{\alpha}{2}}$，则拒绝 H_0。

若$\chi^2_{1-\frac{\alpha}{2}} < \chi^2 < \chi^2_{\frac{\alpha}{2}}$ 则不能拒绝H_0。这种检验法叫χ^2检验法。

例5.3.5 已知维尼纶纤度在正常条件下服从正态分布$N(1.405, (0.048)^2)$，某日抽取 5 根纤维，测得其纤度为1.32，1.55，1.36，1.40，1.44．问这一天纤度的总体标准差是否正常？($\alpha=0.10$)

解　这里要检验的假设为 H_0：$\sigma=\sigma_0=0.048$，计算 χ^2 的观测值得

$$\chi^2=\frac{\sum(x_i-\mu_0)^2}{\sigma_0^2}=13.5$$

自由度为 5，对 $\alpha=0.10$，查表得

$$\chi^2_{\frac{\alpha}{2}}=11.07 \qquad \chi^2_{1-\frac{\alpha}{2}}=1.145$$

由于13.5>11.07，所以拒绝H_0，即认为总体标准差显著地变大了．

(2) 在μ未知时，检验 H_0：$\sigma^2=\sigma_0^2$．

用统计量

$$\chi^2=\sum_{i=1}^{n}\left(\frac{X_i-\overline{X}}{\sigma}\right)^2$$

当 $\sigma=\sigma_0$ 时，它服从自由度为 $n-1$ 的χ^2分布。这时只是χ^2变量的自由度与(1)不同，其他与(1)完全一样．

(3) 检验两个正态总体$X\sim N(\mu_1,\ \sigma_1^2)$，$Y\sim N(\mu_2,\ \sigma_2^2)$的方差是否相等。即检验假设 H_0：$\sigma_1^2=\sigma_2^2$

设 $X_1,\ \cdots, X_{n_1}$ 与 $Y_1,\ \cdots,\ Y_{n_2}$ 分别为由总体X，Y取出的子样，考虑统计量

$$F=\frac{S_1^{*2}/\sigma_1^2}{S_2^{*2}/\sigma_2^2}$$

由于$(n_1-1)S_1^{*2}/\sigma_1^2$，$(n_2-1)S_2^{*2}/\sigma_2^2$分别服从自由度为$n_1-1$，$n_2-1$的$\chi^2$分布，所以$F$服从自由度为$(n_1-1,\ n_2-1)$的$F$分布。当$\sigma_1^2=\sigma_2^2$时，则

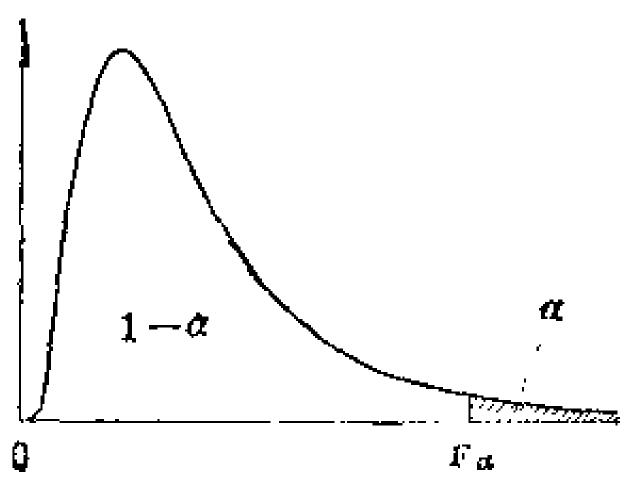
$$F=\frac{S_1^{*2}}{S_2^{*2}}$$

服从自由度为$(n_1-1,\ n_2-1)$的F分布。

对于显著水平α，应求满足下式的$F_{\frac{\alpha}{2}}$，$F_{1-\frac{\alpha}{2}}$

$$P(F\geqslant F_{\frac{\alpha}{2}})=\frac{\alpha}{2}$$

$$P(F\geqslant F_{1-\frac{\alpha}{2}})=1-\frac{\alpha}{2}$$



图（5.3.3）

计算F的观测值，然后根据F值是否在区间$(F_{1-\frac{\alpha}{2}},\ F_{\frac{\alpha}{2}})$之外，决定是否拒绝$H_0$。

附录中的F表，只给出一个上界$P(F\geqslant F_\alpha)=\alpha$，这是因为我们在计算$F$值时，总可以将较大的那一个子样方差作为分子$s_1^{*2}$，所以算得的$F$值总是大于1，因此，只要看它是否大于$F_{\frac{\alpha}{2}}$就行了。

例5.3.6 某一橡胶配方中，原用氧化锌5克，现减为1克，现分别对两种配方各作一批试验，测得橡胶伸长率如下：

用氧化锌5克X：540，533，525，520，545，531，541，529，531，

用氧化锌1克Y：565，577，580，575，556，542，560，532，570，561。

问两种配方的伸长率的总体标准差是否相等。

解 先计算S_x^{*2}，S_y^{*2}的观测值得

$$S_x^{*2}=63.86,\qquad S_y^{*2}=236.8$$

再计算F值得

$$F = 3.7$$

$n_1-1=9$，$n_2-1=8$，对于$\alpha=0.10$，查自由度为（9,8）的F分布表，得$F_{\alpha/2}=F_{0.05}=3.39$。由于$F=3.7>3.39$，所以否定$H_0$，即认为两总体标准差不等（或说有显著差异）。

（三）概率 p 的检验

设某事件A发生的概率p未知，今进行n次独立试验，设事件A发生了k次，以此来检验假设H_0：$p=p_0$。

考虑对应于事件A的随机变量X，$P(X=1)=P(A)=p$，$P(X=0)=P(\overline{A})=1-p$，由于$EX=\mu=p$，所以检验$H_0$：$p=p_0$的问题，就是检验$X$的均值问题。由中心极限定理知统计量

$$\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}=\frac{\frac{k}{n}-p}{\sqrt{p(1-p)/n}}$$

渐近服从正态分布$N(0,1)$，因此当n很大时可用U检验法进行检验。

例5.3.7　从一批产品中随机取出400件，结果有12个次品，试检验假设H_0：$p=0.02$。（$\alpha=0.05$，p为产品次品率）

解　由于当H_0成立时

$$U=\frac{k/n-0.02}{\sqrt{0.02\times0.98/400}}$$

渐近服从正态分布$N(0,1)$。对于$\alpha=0.05$，查正态分布表得$u_\alpha=1.96$，算U的观测值$u=\dfrac{12/400-0.02}{\sqrt{0.02\times0.98/400}}=1.429$ 由于$u<1.96$，所以不能拒绝$p=0.02$。

在上面所有的假设检验问题中，根据实际推断原理，实际上我们都只能作出是否拒绝假设H_0的结论。当不能拒绝H_0时，如果要

想对能否肯定 H_0 作出结论，还必须再作试验。比如在例3中，假若算得 t 值为 2.2，则就不能轻易认为乙地区与甲地区平均最高冻深无显著差异。因为2.2与2.262相差太小了。

§5.4 区间估计

在第二节讲了未知参数的点估计，并指出了某些估计量所具有的优良性质。但是，在许多实际问题中，我们还往往希望通过子样给出一个范围，使得这个范围能够按足够大的给定的概率包含我们所感兴趣的参数。在数理统计学中，这个范围称为置信区间（或置信域），有关的这类问题称为区间估计问题。准确地讲，

设 X_1，X_2，…，X_n 是从总体 X 中取出的子样，选择适当的两个子样函数 $\hat{\theta}_1(X_1, X_2, \cdots, X_n)$，$\hat{\theta}_2(X_1, X_2, \cdots, X_n)$，对于已给的正数 $1-\alpha$（例如0.95，0.90等），若下式成立

$$P[\hat{\theta}_1(X_1, X_2, \cdots, X_n)<\theta<\hat{\theta}_2(X_1, X_2, \cdots, X_n)]=1-\alpha \quad (5.4.1)$$

则称 $(\hat{\theta}_1, \hat{\theta}_2)$ 为 θ 的置信度为 $1-\alpha$ 的置信区间。$\hat{\theta}_1$，$\hat{\theta}_2$ 称为置信界。（5.4.1）式的意思是说，我们能够以 $1-\alpha$ 的可靠程度相信参数在区间 $(\hat{\theta}_1, \hat{\theta}_2)$ 之中。这种用一个区间去估计参数的方法叫做区间估计法。对于区间估计主要解决的问题是，对于已给的置信度 $1-\alpha$，如何求置信区间。下面我们具体加以说明。

（一）正态总体均值 μ 的区间估计

关于正态总体均值 μ 的区间估计，分下面两种情形讨论。

1 当 σ^2 已知时

已知 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$ 是 μ 的无偏、有效、一致估计。由于 X 服从正态分布 $N(\mu, \sigma^2)$，根据定理3.5.2可知 $\sum_{i=1}^{n}X_i$ 服从正态分布

$N(n\mu,\ n\sigma^2)$，所以$\overline{X}\sim N\left(\mu,\ \frac{\sigma^2}{n}\right)$因此，

$$U=\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}$$

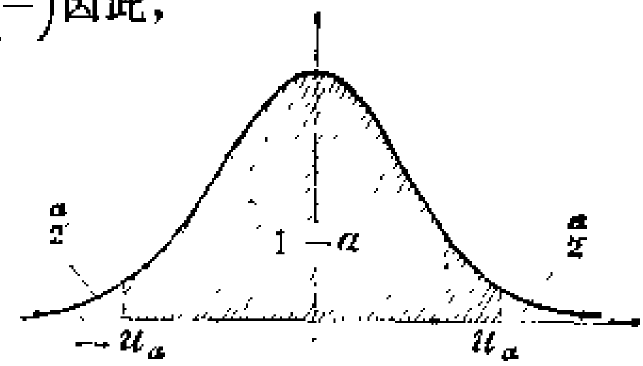


图 (5.4.1)

服从正态分布 $N(0,\ 1)$。对于$1-\alpha$，查正态分布表可得满足下式的u_α

$$1-\alpha=P\left(-u_\alpha<\frac{(\overline{X}-\mu)\sqrt{n}}{\sigma}<u_\alpha\right)$$

$$=P\left(\overline{X}-u_\alpha\frac{\sigma}{\sqrt{n}}<\mu<\overline{X}+u_\alpha\frac{\sigma}{\sqrt{n}}\right)$$

于是得到置信区间 $(\hat{\mu}_1,\ \hat{\mu}_2)$ 为

$$\left(\overline{X}-u_\alpha\frac{\sigma}{\sqrt{n}},\ \overline{X}+u_\alpha\frac{\sigma}{\sqrt{n}}\right) \qquad (5.4.2)$$

结合 1 中求置信区间的过程，给出一般求置信区间的大致步骤如下．

（1）先选择一个由子样 $X_1,\ X_2,\ \cdots,\ X_n$ 及参数 θ 确定的随机变量 $T(X_1,\ X_2,\ \cdots,\ X_n,\ \theta)$。在 1 中就是量$U$，它服从正态分布 N (0, 1) 与μ，σ无关。

（2）求满足

$$P(t_1<T<t_2)=1-\alpha$$

的 $t_1<t_2$，（一般由数表来求），在 1 中就是$-u_\alpha$，u_α。

（3）把括弧内的不等式变形，使被估参数夹在某一区间内

$$P\{\hat{\theta}_1(X_1,X_2,\cdots,X_n)<\theta<\hat{\theta}_2(X_1,X_2,\cdots,X_n)\}=1-\alpha$$

即可求得 $\hat{\theta}_1$，$\hat{\theta}_2$。在 1 中就是

$$\hat{\mu}_1=\overline{X}-u_\alpha\frac{\sigma}{\sqrt{n}},\quad \hat{\mu}_2=\overline{X}+u_\alpha\frac{\sigma}{\sqrt{n}}$$

例5.4.1 从正态总体 $N(\mu,\ 1/4)$ 中取出容量 $n=100$的子样，

算得子样均值为$\overline{X}=13.2$，对置信度0.95，求μ的置信区间。

解 $$\hat{\mu}_1=\overline{X}-1.96\frac{1}{2\sqrt{n}}=13.2-1.96\frac{1}{20}\doteq 13.1$$

$$\hat{\mu}_2=\overline{X}+1.96\frac{1}{2\sqrt{n}}=13.2+1.96\frac{1}{20}\doteq 13.3$$

所以μ的置信度为0.95的置信区间近似为（13.1，13.3）。

根据与上面同样的道理，可以对两个正态总体$X\sim N(\mu_1,\sigma_1^2)$，$Y\sim N(\mu_2,\sigma_2^2)$（$\sigma_1$，$\sigma_2$已知），求总体均值差$\mu_1-\mu_2$的置信度为$1-\alpha$的置信区间。

从总体X与Y中分别取容量为n_1，n_2的子样X_1，$X_2,\cdots,X_{n_1}$，Y_1，$Y_2,\cdots,Y_{n_2}$，由于$\overline{X}=\frac{1}{n_1}\sum_{i=1}^{n_1}X_i$，$\overline{Y}=\frac{1}{n_2}\sum_{i=1}^{n_2}Y_i$分别服从分布$N\left(\mu_1,\frac{\sigma_1^2}{n_1}\right)$，$N\left(\mu_2,\frac{\sigma_2^2}{n_2}\right)$，可知$\overline{X}-\overline{Y}$服从正态分布$N\left(\mu_1-\mu_2,\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}\right)$，因此

$$U=\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$$

服从$N(0,1)$。对于置信度$1-\alpha$，查表可求满足下式的u_α，

$$1-\alpha=P\left(-u_\alpha<\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}<u_\alpha\right)$$

$$=P\left(\overline{X}-\overline{Y}-u_\alpha\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}<\mu_1-\mu_2<\overline{X}-\overline{Y}+u_\alpha\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$

于是得置信区间为

$$\left(\overline{X}-\overline{Y}-u_\alpha\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},\ \overline{X}-\overline{Y}+u_\alpha\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right) \tag{5.4.3}$$

例5.4.2 从两个正态总体$X\sim N(\mu_1,\ 3^2)$，$X\sim N(\mu_2,\ 4^2)$中分别取容量为25，30的子样，算得$\overline{X}=95$，$\overline{Y}=90$，求$\mu_1-\mu_2$的置信度为0.90的置信区间。

解 由（6.3.3）知所求置信区间为

$$\left(95-90-1.65\sqrt{\frac{9}{25}+\frac{16}{30}}\quad 95-90+1.65\sqrt{\frac{9}{25}+\frac{16}{30}}\right)$$

$$\doteq(3.44,\ 6.56)$$

2 当σ^2未知时

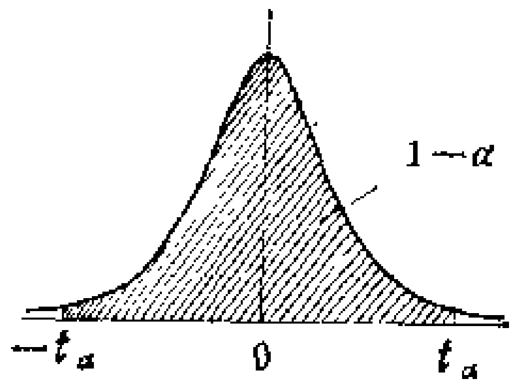
考虑量

$$T=\frac{\overline{X}-\mu}{S}\sqrt{n-1}$$

由定理 5.1.2 推论 1 已知它服从自由度为 $n-1$ 的 t 分布，其中 $S=\sqrt{\frac{1}{n}\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}$ 。对置信度 $1-\alpha$，查 t 分布表，求满足下式的t_α，

$$1-\alpha=P\left(-t_\alpha<\frac{\overline{X}-\mu}{S}\sqrt{n-1}<t_\alpha\right)$$

$$=P\left(\overline{X}-t_\alpha\frac{S}{\sqrt{n-1}}<\mu<\overline{X}+t_\alpha\frac{S}{\sqrt{n-1}}\right)$$



图（5.4.2）

于是得μ的置信区间为

$$\left(\overline{X}-t_\alpha\frac{S}{\sqrt{n-1}},\ \overline{X}+t_\alpha\frac{S}{\sqrt{n-1}}\right)\tag{5.4.4}$$

例 5.4.3 下面是14名青少年足球运动员在比赛前的脉搏，求μ的置信度为0.95的置信区间

11 13 12 12 13 16 11 11

15 12 12 13 11 11

解 首先求$\overline{x}$，s，得$\overline{x}=12.357$，$s=1.493$，$\frac{s}{\sqrt{n-1}}=\frac{s}{\sqrt{13}}\doteq0.414$，对于$1-\alpha=0.95$，查 t 分布表得 $t_\alpha=2.16$，因此所求的

置信区间为

$$(\bar{x}-2.16\times0.414,\ \bar{x}+2.16\times0.414)=(11.46,13.25).$$

下面考虑两个正态总体 $N(\mu_1,\ \sigma_1^2)$，$N(\mu_2,\ \sigma_2^2)$ 的均值差 $\mu_1-\mu_2$的区间估计问题，这里 $\sigma_1=\sigma_2$ 但具体值未知.

设 $X_1,\ \cdots,\ X_{n_1}$，$Y_1,\ \cdots,\ Y_{n_2}$ 分别为取自两总体的子样，则由定理 5.1.2 的推论 3 知，随机变量

$$T=\frac{\bar{X}-\bar{Y}-(\mu_1-\mu_2)}{\sqrt{\dfrac{(n_1S_1^2+n_2S_2^2)(n_1+n_2)}{n_1n_2(n_1+n_2-2)}}}$$

服从自由度为 n_1+n_2-2的 t 分布.

令 $\delta=\sqrt{\dfrac{(n_1+n_2)(n_1S_1^2+n_2S_2^2)}{n_1n_2(n_1+n_2-2)}}$，则所求的置信区间为

$$(\bar{X}-\bar{Y}-t_\alpha\cdot\delta,\ \bar{X}-\bar{Y}+t_\alpha\cdot\delta) \tag{5.4.5}$$

（二）正态总体方差 σ^2 的区间估计

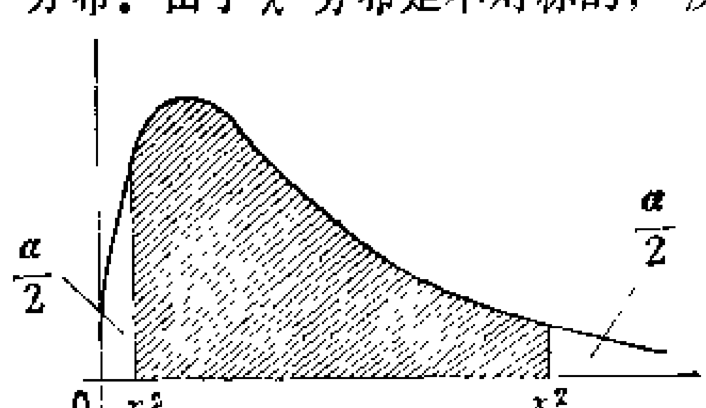
与求 μ 的置信区间的方法类似，我们分总体均值 μ 已知与未知两种情形，求 σ^2 的置信区间.

1　当 μ 已知时

设 $X_1,\ X_2,\ \cdots,\ X_n$ 为取自正态总体 $N(\mu,\ \sigma^2)$ 的子样. 则 $(X_i-\mu)/\sigma, i=1,\ 2,\ \cdots,\ n$，服从正态分布$N(0,1)$. 所以$\sum\limits_{i=1}^{n}(X_i-\mu)^2/\sigma^2$服从自由度为$n$ 的 χ^2 分布. 由于 χ^2 分布是不对称的，所以对于 $1-\alpha$，查 χ^2 分布表时应求满足下式的 $\chi_1^2<\chi_2^2$

$$P(\chi^2\geqslant\chi_2^2)=\frac{\alpha}{2},$$

$$P(\chi^2\geqslant\chi_1^2)=1-\frac{\alpha}{2}，则$$



$$1-\alpha=P\left(\chi_1^2<\sum_{i=1}^{n}(X_i-\mu)^2/\sigma^2<\chi_2^2\right)$$

图 (5.4.3)

$$=P\left(\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi_2^2}<\sigma^2<\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi_1^2}\right)$$

因此得所求置信区间为

$$\left(\sum_{i=1}^{n}(X_i-\mu)^2/\chi_2^2,\quad \sum_{i=1}^{n}(X_i-\mu)^2/\chi_1^2\right) \tag{5.4.6}$$

例5.4.4　下面是取自正态总体 $N(0,\sigma^2)$ 的子样，求 σ^2 的置信度为0.98的置信区间，

1.30，－2.78，0.06，1.98

解　查自由度为 4 的 χ^2 分布表。

$P(\chi^2>13.28)=0.01$

$P(\chi^2>0.297)=0.99$

所以 $\chi_1^2=0.297$，$\chi_2^2=12.28$，$\sum_{i=1}^{4}x_i^2=13.3424$ 代入(5.4.6) 得置信区间为

$(13.3424/13.28,\ 13.3424/0.297)\doteq(1.005,\ 44.924)$

2　当 μ 未知时

设 X_1，X_2，…，X_n 为取自正态总体 $N(\mu,\sigma^2)$ 的子样，则

$$\chi^2=\frac{nS^2}{\sigma^2}$$

服从自由度为 $n-1$ 的 χ^2 分布。求满足

$$1-\alpha=P\left(\chi_1^2<\frac{nS^2}{\sigma^2}<\chi_2^2\right)=P\left(\frac{nS^2}{\chi_2^2}<\sigma^2<\frac{nS^2}{\chi_1^2}\right)$$

的 χ_1^2，χ_2^2。则置信度为 $1-\alpha$ 的置信区间为

$$\left(\frac{nS^2}{\chi_2^2},\ \frac{nS^2}{\chi_1^2}\right) \tag{5.4.7}$$

例 5.4.5 对例 5.4.3 的数据，求 σ^2 的置信度为0.90的置信区间。

解　已算得 $S^2=2.11$，查自由度为 13 的 χ^2 分布表得 $P(\chi^2>22.362)=0.05$，$P(\chi^2>5.892)=0.95$。所以 $\chi_1^2\doteq5.892$，$\chi_2^2\doteq$

22.362，将S^2，χ_1^2，χ_2^2代入（5.4.7）得置信区间为

$$(14\times2.11/22.362,\ 14\times2.11/5.892)\doteq(1.3,\ 5)$$

（三）百分率 p 的区间估计

设某事件 A 发生的概率为 p，求 p 的置信度为 $1-\alpha$ 的置信区间。

考虑每次试验中 A 出现的次数 X，它服从贝努利分布，$P(X=1)=P(A)=p$，$P(X=0)=P(\overline{A})=1-p=q$，由于 $EX=\mu=p$，所以求 p 的置信区间就是求总体 X 的均值 μ 的置信区间，因此，容易想到用随机变量

$$\frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}}=\frac{\overline{X}-p}{\sqrt{p(1-p)/n}}$$

这时由于 X 不是服从正态分布，所以这个随机变量不服从正态 $N(0,1)$。但是，根据中心极限定理 $P\left(\frac{\sum\limits_{i=1}^{n}X_i-n\mu}{\sigma\sqrt{n}}\leqslant x\right)\to\frac{1}{\sqrt{2\pi}}\int^{x}e^{-\frac{y^2}{2}}dy$ 它渐近服从正态 $N(0,1)$，所以当 n 充分大时，仍可用（一）中（1）的方法做区间估计，即对 $1-\alpha$，查正态分布表，求满足下式的 u_α

$$1-\alpha=P\left(-u_\alpha<\frac{\overline{X}-p}{\sqrt{p(1-p)/n}}<u_\alpha\right)$$

$$=P(\overline{X}-u_\alpha\sqrt{p(1-p)/n}<p<\overline{X}+u_\alpha\sqrt{p(1-p)/n})$$

在不等式两端仍含有未知量 p，由于 n 充分大，通常我们可用频率 $\frac{k}{n}=\overline{X}$ 代替两端的 p，这样便得到置信区间为

$$(\overline{X}-u_\alpha\sqrt{\overline{X}(1-\overline{X})/n},\ \overline{X}+u_\alpha\sqrt{\overline{X}(1-\overline{X})/n})$$

不过这时只能说置信度约为 $1-\alpha$，因为由于用 $\overline{X}$ 代替 p 因而又产生

一次误差。

可以看出，假设检验问题与区间估计问题的关系是很明显的。若参数 θ 的一个置信度为 $1-\alpha$ 的置信区间为 $(\hat{\theta}-\lambda, \hat{\theta}+\lambda)$，则对假设 H_0：$\theta=\theta_0$ 的检验法就是：若 θ_0 在区间 $(\hat{\theta}-\lambda, \hat{\theta}+\lambda)$ 外就拒绝 H_0，否则不能拒绝 H_0，这时显著水平为 α。

§5.3，§5.4两节的主要结果可概括成表（见下页）。

§5.5 非参数检验

第三节中讲的各种参数检验，几乎都是在正态总体的假定下进行的，但是，在很多实际问题中，对于总体是否服从正态分布并不知道，这时，需要根据子样对总体服从正态分布的假设进行检验。

在总体 X 是连续型随机变量的情形，首先我们应该根据子样的观测值，画出直方图，以期对总体所服从的分布有一个大致的了解。直方图的作法如下：

设 n 次观测数据为 $x_1, x_2, \cdots, x_n$。先找出其中最小值 a 和最大值 b，然后将区间 $[a, b]$ 或比它略大一点的区间分成若干个相等的小区间 $[a,a_1), [a_1,a_2), \cdots, [a_{l-1}, b)$，然后计算落在各小区间里的数据的个数及数据的频率，这样，把全部数据分成了 l 组每组有相应的频率，在横轴上，以各小区间长为底，以相应的 $\dfrac{\text{频率}}{\text{小区间长}}$ 为高，画小矩形，也就是使小矩形面积等于频率。这样就得到了由许多小矩形构成的直方图。

例5.5.1 下面是某地50年的四月份的平均气温的资料：

4.0	6.6	5.2	6.4	7.9	8.6	3.0	4.4	6.7	6.9
4.7	9.0	6.8	8.6	5.2	5.8	7.9	5.6	8.8	7.8
5.7	8.4	4.1	6.4	6.2	5.2	6.8	5.6	5.6	8.1
8.2	6.4	4.8	6.9	7.0	9.7	6.4	7.3	6.8	6.8
4.8	5.8	6.5	5.9	7.3	5.5	7.4	6.2	7.7	7.1

表5.4.1 关于正态总体参数的显著性检验一览表

	假设 H_0		统计量	拒绝域	分布表	自由度	名称
单子样	$\mu=\mu_0$	$\sigma=\sigma_0$ 已知	$U=\dfrac{\overline{X}-\mu_0}{\dfrac{\sigma_0}{\sqrt{n}}}$	$\lvert u\rvert \geq u_\alpha$	正态表		U 检验
		σ 未知	$T=\dfrac{\overline{X}-\mu_0}{\dfrac{S}{\sqrt{n-1}}}$	$\lvert t\rvert \geq t_\alpha$	t 分布表	$n-1$	t 检验
	$\sigma=\sigma_0$	$\mu=\mu_0$ 未知	$\chi^2=\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu_0)^2}{\sigma_0^2}$	$\chi^2 \geq \chi^2_{\frac{\alpha}{2}}$ 或 $\chi^2 \leq \chi^2_{1-\frac{\alpha}{2}}$	χ^2 分布表	n	χ^2 检验
		μ 未知	$\chi^2=\dfrac{\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}{\sigma_0^2}$	$\chi^2 \geq \chi^2_{\frac{\alpha}{2}}$ 或 $\chi^2 \leq \chi^2_{1-\frac{\alpha}{2}}$	χ^2 分布表	$n-1$	
双子样	$\mu_1=\mu_2$（σ_1，σ_2 已知）		$U=\dfrac{\overline{X}-\overline{Y}}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}$	$\lvert u\rvert \geq u_\alpha$	正态表		U 检验
	$\mu_1=\mu_2$（$\sigma_1=\sigma_2$ 未知）		$T=\dfrac{\overline{X}-\overline{Y}}{\sqrt{\dfrac{n_1S_1^2+n_2S_2^2}{n_1+n_2-2}\cdot\dfrac{n_2+n_2}{n_1n_2}}}$	$\lvert t\rvert \geq t_\alpha$	t 分布表	n_1+n_2-2	t 检验
	$\sigma_1^2=\sigma_2^2$		$F=\dfrac{S_1^{*2}}{S_2^{*2}}$	$F \leq F_{1-\frac{\alpha}{2}}$ 或 $F \geq F_{\frac{\alpha}{2}}$	F 表	(n_1-1, n_2-1)	F 检验

表5.4.2　关于正态总体参数的区间估计的一览表

项目		条件	随机变量及其分布	置信区间
单个正态总体	关于 μ 的置信区间	σ 已知	$U=\dfrac{\overline{X}-\mu}{\dfrac{\sigma}{\sqrt{n}}}\sim N(0,1)$	$\left(\overline{X}-u_{\alpha}\dfrac{\sigma}{\sqrt{n}},\ \overline{X}+u_{\alpha}\dfrac{\sigma}{\sqrt{n}}\right)$
		σ 未知	$T=\dfrac{\overline{X}-\mu}{\dfrac{S}{\sqrt{n-1}}}\sim t(n-1)$	$\left(\overline{X}-t_{\alpha}\dfrac{S}{\sqrt{n-1}},\ \overline{X}+t_{\alpha}\dfrac{S}{\sqrt{n-1}}\right)$
	关于 σ 的置信区间	μ 已知	$\chi^2(n)=\dfrac{\sum\limits_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}\sim\chi^2(n)$	$\left(\dfrac{1}{\chi_2^2}\Sigma(X_i-\mu)^2,\ \dfrac{1}{\chi_1^2}\Sigma(X_i-\mu)^2\right)$
		μ 未知	$\chi^2(n-1)=\dfrac{\sum\limits_{i=1}^{n}(X_i-\overline{X})^2}{\sigma^2}\sim\chi^2(n-1)$	$\left(\dfrac{1}{\chi_2^2}\Sigma(X_i-\overline{X})^2,\ \dfrac{1}{\chi_1^2}\Sigma(X_i-\overline{X})^2\right)$
两个正态总体	关于 $\mu_1-\mu_2$ 的置信区间	σ_1,σ_2 已知	$U=\dfrac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}}\sim N(0,1)$	$\left(\overline{X}-\overline{Y}-u_{\alpha}\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}},\ \overline{X}-\overline{Y}+u_{\alpha}\sqrt{\dfrac{\sigma_1^2}{n_1}+\dfrac{\sigma_2^2}{n_2}}\right)$
		σ_1,σ_2 未知 $\sigma_1=\sigma_2$	$T=\dfrac{\dfrac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\dfrac{1}{n_1}+\dfrac{1}{n_2}}}}{\sqrt{\dfrac{n_1S_1^2+n_2S_2^2}{n_1+n_2-2}}}\sim t(n_1+n_2-2)$	$\left(\overline{X}-\overline{Y}-t_{\alpha}\sqrt{\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)\dfrac{n_1S_1^2+n_2S_2^2}{n_1+n_2-2}},\ \overline{X}-\overline{Y}+t_{\alpha}\sqrt{\left(\dfrac{1}{n_1}+\dfrac{1}{n_2}\right)\dfrac{n_1S_1^2+n_2S_2^2}{n_1+n_2-2}}\right)$

其中最小值为 3.0，最大值为 9.7，将区间[3，10)分成 7 个小区间，区间长为 1，计算数据落在各小区间里的频数和频率得

温度	频数	频率
[3, 4)	1	1/50
[4, 5)	6	6/50
[5, 6)	11	11/50
[6, 7)	15	15/50
[7, 8)	9	9/50
[8, 9)	6	6/50
[9,10)	2	2/50

对应的直方图为

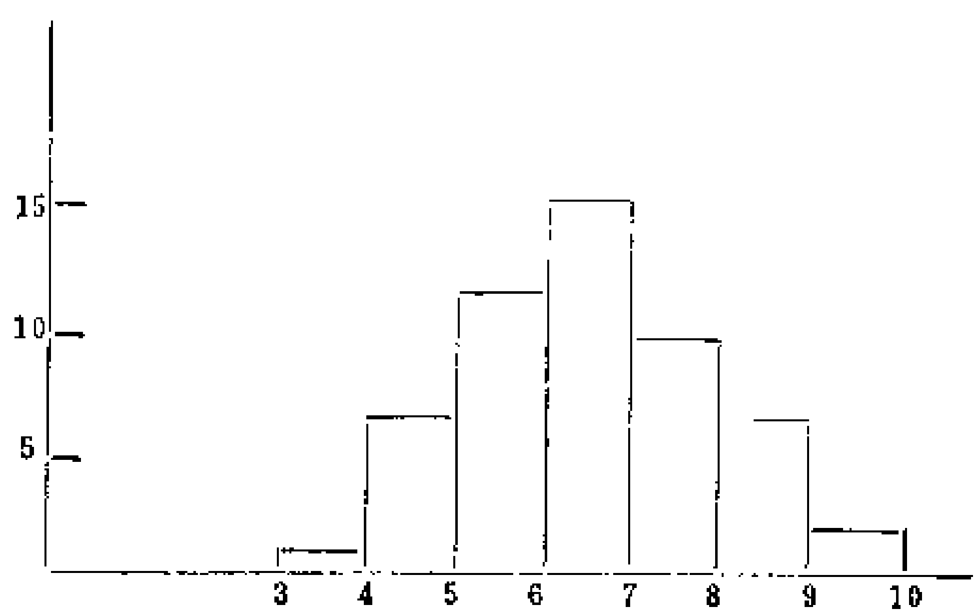


图 (5.5.1)

可以设想，如果子样的容量增大，并且分组更细，则各组出现的频率将稳定在一个确定的数值附近。由此，直方图的形状将变为一条光滑的曲线。由图 (5.5.1) 可以设想，月平均气温将服从正态分布$N(\mu, \sigma^2)$，其中参数μ，σ^2可以分别由子样均值$\overline{x}=6.53$，子样方差$S^2=1.4^2$加以估计。

这样一来，我们可以提出假设 H_0: $F(x)=F_0(x)$，其中 $F_0(x)$ 是正态分布 $N(6.53, 1.4^2)$，尔后根据子样进行显著性检验.

下面我们先考虑一般的情况，尔后再回到我们这个具体问题上.

假设 H_0: $F(x)=F_0(x)$，这里 $F_0(x)$ 是某一已知的分布，要求根据子样进行显著性检验.

对于总体分布的检验法也和对参数的检验法一样，首先要找一个合适的统计量，它能反映实际观测数据与所假设的总体分布之间的差异，并且它在原假设成立时分布已知. 然后对已给的显著水平 α，找出拒绝域，再进行判断.

这里只介绍一个最常用的，用 χ^2 分布进行总体分布检验的方法，称为皮尔逊 (K.pearson)χ^2 检验.

首先根据子样值 $x_1, x_2, \cdots, x_n$ 的范围，把 $(-\infty, +\infty)$ 分成若干个区间 $(-\infty, a_1)$, (a_1, a_2), $\cdots(a_{m-1}, \infty)$，然后记 $x_1, x_2, \cdots, x_n$ 中落在各区间 (a_{i-1}, a_i) 中的个数 ν_i，并且算出当 H_0 成立时 X 落在 (a_{i-1}, a_i) 中的概率 $p_i=P(a_{i-1}<X\leqslant a_i)$，则 np_i 就是理论上 $x_1, \cdots, x_n$ 应落在 (a_{i-1}, a_i) 中的个数.

对每个区间求差 ν_i-np_i，然后平方求和

$$\eta=\sum_{i=1}^{m}\frac{(\nu_i-np_i)^2}{np_i}$$

这个统计量反映了子样的经验分布与 H_0 成立时的理论分布之间的差异. 因此我们可以根据 η 的观测值的大小来推断 H_0 是否正确. 关于 η 的分布有下面的定理.

皮尔逊定理　当 $n\to\infty$ 时，不论假设的总体分布 $F_0(x)$ 是什么分布，这里 $F_0(x)$ 不带有未知参数，则变量 η 的极限分布是自由度为 $m-1$ 的 χ^2 分布，即

$$\lim_{n\to\infty}P(\eta\leqslant x)=P(\chi^2(m-1)\leqslant x)$$

证明　我们对 $m=2$ 的特殊情形给出证明. 对 $m>2$ 的情形的证明，有兴趣的读者可查阅其它专著.

$$\eta = \frac{(\nu_1 - np_1)^2}{np_1} + \frac{(\nu_2 - np_2)^2}{np_2}$$

$$= \frac{(\nu_1 - np_1)^2}{np_1} + \frac{[(n-\nu_1) - n(1-p_1)]^2}{n(1-p_1)}$$

$$= \frac{(\nu_1 - np_1)^2}{np_1} + \frac{(\nu_1 - np_1)^2}{n(1-p_1)}$$

$$= \frac{(\nu_1 - np_1)^2}{np_1(1-p_1)}$$

引进记号

$$\nu_1 = \sum_{k=1}^{n} \zeta_k$$

其中 $\zeta_1, \zeta_2, \cdots, \zeta_n$ 是 n 个相互独立的随机变量，定义如下：

$$\zeta_k = \begin{cases} 1, & X_k \in [a_0, a_1] \\ 0 & X_k \overline{\in} [a_0, a_1], \ k = 1, 2, \cdots, n \end{cases}$$

则

$$P(\zeta_k = 1) = F_0(a_1) - F_0(a_0) = p_1$$

$$E\nu_1 = np_1, \quad D\nu_1 = np_1(1-p_1)$$

由第四章的定理 4.3.1 知，统计量

$$\frac{\nu_1 - np_1}{\sqrt{np_1(1-p_1)}}$$

的极限分布为 $N(0, 1)$，所以统计量 η 的极限分布为 $\chi^2(1)$。

因此统计量 η 可用 χ^2 表示，对于已给的 α，求满足下式的 χ^2_α

$$P(\chi^2 \geqslant \chi^2_\alpha) = \alpha$$

若观测值 $\chi^2 \geqslant \chi^2_\alpha$ 则拒绝 H_0，若 $\chi^2 < \chi^2_\alpha$ 则不能拒绝 H_0。

在上面的讨论中，我们假定 $F_0(x)$ 是完全确定的，因此可以求 p_i，但是在实际中，有时原假设只给出分布的类型，而其参数未知，比如所提假设 H_0 为"总体服从正态分布"，假设 H_0 中没有给出 μ 和 σ^2。这样，上面的方法就不能用了。对于这种情况自然会想到用总体参数的最大似然估计去代替总体参数。可以证明（从略）这时 η 仍渐近服从 χ^2 分布，不过自由度减少了，所减少的自由度的

数目等于被估计的参数的个数．例如对正态总体用$\overline{X}$和S^2代替μ和σ^2以后，变量η渐近服从自由度为$m-3$的χ^2分布．

例 5.5.2　对例5.5.1的数据，假设H_0：$F(x)=N(6.53,1.4^2)$，用皮尔逊χ^2检验法检验，取$\alpha=0.05$。

解　此题主要的工作在于由样本值去算统计量

$$\eta=\sum_{i=1}^{m}\frac{(\nu_i-np_i)^2}{np_i}$$

的值．

取 $a_1=4$，$a_2=5$，$a_3=6$，$a_4=7$，$a_5=8$，$a_6=9$，这些分点把$(-\infty,\ \infty)$分成了七段．当H_0成立时，计算p_i的数值，

$$p_1=F(a_1)$$
$$p_2=F(a_2)-F(a_1)$$
$$p_3=F(a_3)-F(a_2)$$
$$p_4=F(a_4)-F(a_3)$$
$$p_5=F(a_5)-F(a_4)$$
$$p_6=F(a_6)-F(a_5)$$
$$p_7=1-F(a_6)$$

为了算出$F(a_i)$的数值，我们利用标准正态分布函数$\Phi(x)$的数值表．

由于

$$F(a_i)=P(\eta\leqslant a_i)=P\left(\frac{\eta-6.53}{1.4}\leqslant\frac{a_i-6.53}{1.4}\right)$$
$$=\Phi\left(\frac{a_i-6.53}{1.4}\right)$$

因而

$$F(a_1)=\Phi\left(\frac{4-6.53}{1.4}\right)=\Phi(-1.81)=1-\Phi(1.81)$$
$$=0.0351$$

$$F(a_2)=\Phi\left(\frac{5-6.53}{1.4}\right)=\Phi(-1.09)=1-\Phi(1.09)$$
$$=0.1379$$

$$F(a_3)=\Phi\left(\frac{6-6.53}{1.4}\right)=\Phi(-0.38)=1-\Phi(0.38)$$

$$=0.3520$$

$$F(a_4)=\Phi\left(\frac{7-6.53}{1.4}\right)=\Phi(0.34)=0.6331$$

$$F(a_5)=\Phi\left(\frac{8-6.53}{1.4}\right)=\Phi(1.05)=0.8531$$

$$F(a_6)=\Phi\left(\frac{9-6.53}{1.4}\right)=\Phi(1.76)=0.9608$$

由此可算出

$p_1=0.0351$, $p_2=0.1028$,

$p_3=0.2141$, $p_4=0.2811$,

$p_5=0.2200$, $p_6=0.1077$,

$p_7=0.0392$.

计算 η 的数值时，可列出下表

区间	$(-\infty,4)$	$[4,5)$	$[5,6)$	$[6,7)$	$[7,8)$	$[8,9)$	$[9,+\infty)$
ν_i	1	6	11	15	9	6	2
p_i	0.0351	0.1028	0.2141	0.2811	0.2200	0.1077	0.0392
np_i	1.755	5.14	10.705	14.055	11	5.385	1.96
$(\nu_i-np_i)^2$	0.5700	0.7396	0.0870	0.8930	4	0.3782	0.0016
$\frac{(\nu_i-np_i)^2}{np_i}$	0.3248	0.1439	0.0081	0.0635	0.3636	0.0702	0.0008

$$\eta=\sum_{i=1}^{7}\frac{(\nu_i-np_i)^2}{np_i}=0.9749$$

查 χ^2 分布表，$\alpha=0.05$，自由度为$7-1-2=4$，得 $\chi^2_{0.05}=9.488$，现在 $0.9749<9.488$，故不拒绝 H_0。

当我们用χ^2 检验法时，必须注意它是建立在极限定理的基础上的，因此子样容量不能太小，一般最好不小于50。

皮尔逊χ^2检验有广泛的应用．它的优点是，不管$F_0(x)$是怎样的分布函数，都可以检验一个总体是否以它为分布函数．对于连续型变量的情形，它的计算量较大，但如果使用电子计算器，则这点计算量就不算什么了．对于离散型变量的情形，手算也不费事．

例5.5.3 在一个正20面体的20个面上，分别标以数字0，1，2，…，9，每个数字在两个面上标出，为检验它的均匀性，共作800次投掷试验，数字 0，1，…，9 朝正上方的次数列成下表

数　字	0	1	2	3	4	5	6	7	8	9
频　数	74	92	83	79	80	73	77	75	76	91

问该正20面体是否匀称？

解　假设该正20面体是匀称的，即X的分布为

$$p_i = P(X=i) = \frac{1}{10} \qquad i=0,\ 1,\ 2,\ \cdots,9$$

$$np_i = 800 \times \frac{1}{10} = 80$$

为计算η的数值，将主要计算结果列出

数　　字	0	1	2	3	4	5	6	7	8	9
频　数 ν_i	74	92	83	79	80	73	77	75	76	91
$\nu_i - np_i$	−6	12	3	−1	0	−7	−3	−5	−4	11
$(\nu_i - np_i)^2$	36	144	9	1	0	49	9	25	16	121

由此可算出

$$\chi^2 = \sum_{i=0}^{9} \frac{(\nu_i - np_i)^2}{np_i} = 5.125$$

查χ^2分布表，自由度为$10-1=9$，$\alpha=0.05$，得

$$\chi^2_{0.05} = 16.9$$

现在 $5.125<16.9$ 故不能拒绝H_0，即没有发现这个正20面体不匀称．

习　　题

1．若总体X服从参数为p的两点分布，即　$P(X=1)=p$，$P(X=0)=q$，$(p+q=1)$，设(X_1, X_2, X_3)是由X中抽取的一个简单随机子样，

（1）写出它的子样空间与(X_1, X_2, X_3)的概率分布；

（2）如果$(0,1,1)$是(X_1, X_2, X_3)的一个观测值，计算它的子样均值与子样方差．

2．设总体X服从正态$N(\mu, \sigma^2)$分布，$(X_1, X_2, \cdots, X_5)$是从X中抽取的一个简单随机子样，

（1）写出它的子样空间与$(X_1, X_2, \cdots, X_5)$的分布密度；

（2）如果$(0, 1, -1, 2, -2)$是$(X_1, X_2, \cdots, X_5)$的一个观测值，计算它的子样均值与子样方差．

3．设$X_1, X_2, \cdots, X_n$是取自任意总体X的子样，令$X'_i=X_i+a$，a为任一常数．证明

$$\overline{X}'=\frac{1}{n}\sum_{i=1}^{n} X'_i=\overline{X}+a$$

$$S'^2=\frac{1}{n}\sum_{i=1}^{n}(X'_i-\overline{X}')^2=S^2$$

在计算$\overline{X}$和S^2时，如果数据都比较大，可共同减去一个常数，然后按上面的关系式进行计算．对下列一组数据求其均值与方差：

33，36，34，36，36，35，31，35，33，27

4．设总体X服从二项分布$b(n, p)$，n为正整数，$0<p<1$，它们均为未知参数．$X_1, X_2, \cdots, X_n$是从X中抽取的大小为n的子样，试求n与p的矩法估计．

5．设总体X服从参数为λ的普阿松分布，$\lambda>0$为未知参数．$X_1, X_2, \cdots, X_n$是从X中抽取的大小为n的子样，试求λ的矩法

估计与最大似然估计。

6. 设总体X具有如下分布，求参数的最大似然估计

(1) $p(x;\ \theta_1,\ \theta_2)=\begin{cases}\dfrac{1}{\theta_2-\theta_1} & \theta_1\leqslant x_2\leqslant\theta_2\\ 0 & 其它\end{cases}$

(2) $p(x;\ \theta)=\begin{cases}\dfrac{1}{\theta}e^{-\frac{x}{\theta}} & (x\geqslant0,\ \theta>0)\\ 0 & (x<0)\end{cases}$

7. 设X服从两点分布$\begin{pmatrix}0, & 1\\ 1-p, & p\end{pmatrix}$，即

$P(X=k)=p^k(1-p)^{1-k}\qquad k=0,\ 1$

求p的最大似然估计。

8. 设 X_1，X_2 为来自总体X的容量为2的子样，X具有如下分布

$$p(x;\theta)=\begin{cases}\dfrac{2(\theta-x)}{\theta^2} & (0<x<\theta);\\ 0 & 其它\end{cases}$$

(1) 求θ的最大似然估计

(2) 求总体均值μ的最大似然估计

9. 设总体X服从均匀分布

$$p(x;\theta)=\begin{cases}\dfrac{1}{\theta} & (0<x\leqslant\theta)\\ 0 & 其它\end{cases}$$

证明$\hat{\theta}=2\overline{X}$是$\theta$的无偏估计。

10. 设X_1，X_2，…，X_n是取自均值与方差分别为μ与σ^2的总体X的子样，取$\hat{\mu}=c_1x_1+c_2x_2+\cdots+c_nx_n$作为总体均值$\mu$的估计量，问$c_i$为什么值时，$\hat{\mu}$是无偏的，且$\hat{\mu}$的方差最小。

11. 证明子样均值的平方$\overline{X}^2$不是总体均值平方μ^2的无偏估计。

（可见若$\hat{\theta}$是θ的无偏估计，并不能推出$f(\hat{\theta})$是$f(\theta)$的无偏估计。）

12．证明当总体X的均值μ已知时，则

$$S_0^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$$

是总体方差σ^2的无偏估计。

13．证明总体均值μ的无偏估计$\hat{\mu}=\sum_{i=1}^{n}c_iX_i\left(c_i\geqslant 0,\sum_{i=1}^{n}c_i=1\right)$，当$\lim\limits_{n\to\infty}\sum_{i=1}^{n}c_i^2=0$时，为$\mu$的一致估计。

14．设$\hat{\theta}_1$，$\hat{\theta}_2$是θ的两个独立的无偏估计，并且$\hat{\theta}_1$的方差是$\hat{\theta}_2$的方差的2倍，试找出常数k_1，k_2使得$k_1\hat{\theta}_1+k_2\hat{\theta}$是$\theta$的无偏估计，并且在所有这样的线性估计量中方差最小。

15．设X_1，X_2，X_3是来自总体X的子样，X的分布密度为

$$f(x;\theta)=\begin{cases}\dfrac{1}{\theta} & 0<x<\theta \\ 0 & \text{其它}\end{cases}\qquad 0<\theta<\infty$$

试证$\frac{4}{3}\max\limits_{1\leqslant i\leqslant 3}X_i$、$4\min\limits_{1\leqslant i\leqslant 3}X_i$都是$\theta$的无偏估计量，并比较它们哪个有效。

16．设X_1，X_2，…，X_n是来自正态总体$N(\theta,1)$（其中$0<\theta\leqslant 1/10$）的子样，对于θ的两个估计量$\hat{\theta}_1=\overline{X}$，$\hat{\theta}_2=0.05$，比较它们与$\theta$的误差的平方的均值

$$E[(\hat{\theta}_1-\theta)^2],\ E[(\hat{\theta}_2-\theta)^2].$$

17．某种产品的单个重量的均值是$12g$，标准差是$1g$．更新设备以后，从所生产产品中随机取出100个，测得子样均值$\overline{x}=12.5g$，问设备更新前后产品的平均重量是否有变化．（$\alpha=0.10$）

18．从正态总体中随机取出容量为20的子样，算得$\overline{x}=42$．$s=5$．要求检验假设H_0：$\mu=50$（$\alpha=0.01$）。

19. 对10名患者试用两种药品A，B，各药用后，分别测延长睡眠时间，结果得$\overline{x}_A=2.33$，$s_A^2=3.61$　$\overline{x}_B=0.75$，$s_B^2=2.89$

可否认为A，B对延长睡眠时间的效果没有差异。假设延长睡眠时间服从正态分布，而两者方差没有差别。$(\alpha=0.01)$

20. 已知健康人的红血球直径服从均值为7.20μ的正态分布。今在某一患者血液中随机测得9个红血球的直径如下：（单位μ）

7.1，7.3，7.7，7.8，8.0，8.1，8.5，9.0，7.6。问该患者红血球平均直径与健康人有无显著差异？$(\alpha=0.05)$

21. 某种物品在处理前后，分别取子样分析其含脂率如下：

处理前：0.19　0.18　0.21　0.30　0.66　0.42　0.08　0.12　0.30　0.27

处理后：0.15　0.13　0.00　0.07　0.24　0.24　0.19　0.04　0.08　0.20　0.12

假定处理前后含脂率服从正态分布，且标准差不变，问处理前后含脂率的总体均值有无显著变化？

22. 在正态总体$N(\mu,\sigma^2)$（σ已知）的均值$\mu=\mu_0$的检验中，设$\overline{x}$为容量为n的子样均值，若$|\overline{x}-\mu_0|\geqslant k\sigma$，则否定$\mu=\mu_0$。这时为使显著水平在0.05以下，$n$至少应为多少？对$k=1$，$1/2$，$1/3$三种情况进行考虑。

23. 从正态总体$N(\mu,\sigma^2)$（μ未知）中随机取出容量为8的子样，得$\overline{x}=61$，$\sum\limits_{i=1}^{8}(x_i-61)^2=652.8$，试以显著水平0.05检验假设$H_0:\sigma^2=8^2$。

24. 从正态总体$N(10,\sigma^2)$中随机取容量为5的子样为9.4，11.3，8.7，10.6，9.7，试以显著水平$\alpha=0.05$，检验假设H_0：$\sigma=1$。

25. 从两个正态总体X，Y中分别取容量为9，11的子样，算得$\sum\limits_{i=1}^{9}(x_i-\overline{x})^2=96$，$\sum\limits_{i=1}^{11}(y_i-\overline{y})^2=45$。试以显著水平$\alpha=0.02$检

验两个总体的方差是否相等。

26. 调查了144人的每日吸烟量，得 $\overline{X}=12$（支），假定吸烟量服从正态分布$N(\mu,\sigma^2)$，已知 $\sigma=4$，求 μ 的置信区间，（1）置信度为0.95，（2）置信度为0.99。

27. 设总体X服从正态分布$N(\mu,\sigma^2)$，σ^2未知，对于以下已知值，求总体均值μ的置信区间（置信度为0.95）。

（1）$n=5$，$\overline{x}=12.3$，$s=0.2$

（2）$n=21$，$\overline{x}=13.2$，$s^2=5$

（3）$n=36$，$\overline{x}=15.2$，$s^{*2}=68.4$

28. 设总体X服从正态分布 $N(\mu,\sigma^2)$，μ未知，对于以下已知值，求总体方差σ^2的置信区间（置信度为$1-\alpha$）。

（1）$n=25$　$\sum_{i=1}^{n}x_i=101.1$　$\sum_{i=1}^{n}x_i^2=412.75$，$\alpha=0.10$。

（2）$n=12$　$s^2=1.243$　　$\alpha=0.02$

29. 设总体X服从正态分布 $N(\mu,10^2)$，若使μ的置信度为0.95的置信区间长为5，试问子样容量n最小应为多少？又置信度为0.99时n应为多少？

30. 下面是取自正态总体X的子样

50.7	69.8	54.9	53.4	54.3
66.1	44.8	48.1	42.2	35.7

求标准差σ的置信度为0.90的置信区间。

31. 从两个正态总体 $N(\mu,50)$，$N(\mu',60)$ 分别取出容量为10，30的子样，算得它们的子样均值为$\overline{x}=80$，$\overline{y}=70$。求 $\mu-\mu'$的置信度为0.95的置信区间。

32. 随机取100个产品，其中一等品为64个，求一等品的概率p的置信度为0.95的置信区间。

33. 设$X_1,\cdots X_{n_1}$；$Y_1,\cdots,Y_{n_2}$为取自两个正态总体 $N(\mu_1,\sigma_1^2)$，$N(\mu_2,\sigma_2^2)$ 的子样，求 $\mu_1-\mu_2$的置信度为0.90的置信区间。（σ_1^2，σ_2^2为已知）

34. 某食品厂为加强质量管理，对某天生产的罐头抽查了 100 个（数据见下表），用皮尔逊 χ^2 检验看它是否服从正态分布？（$\alpha=0.10$）

100个罐头样品的净重数据（单位：克）

342	340	348	346	343	342	346	341	344	348	346	346
340	344	342	344	345	340	344	344	343	344	342	343
345	339	350	337	345	349	336	348	344	345	332	342
342	340	350	343	347	340	344	353	340	340	356	346
345	346	340	339	342	352	342	350	348	344	350	335
340	338	345	345	349	336	342	338	343	343	341	347
341	347	344	339	347	348	343	347	346	344	345	350
341	338	343	339	343	346	342	339	343	350	341	346
341	345	344	342								

35. 有一正四面体，将此四面体的四面分别涂为红、黄、蓝、白四种不同的颜色，现作如下的抛掷试验：任意地抛掷该四面体，直到白色的一面与地面相接触为止，记录下抛掷的次数. 作这样的试验 200 次，其结果如下：

抛掷次数	1	2	3	4	$\geqslant 5$
频　　数	56	48	32	28	36

问，该四面体是否均匀？（$\alpha=0.05$）

36. 在某公路上观察汽车通过的频繁情况. 取15秒为一个时间单位，记下过路汽车的辆数，连续观察 200 个时间单位（即共观察50分钟），得数据如下：

过路的辆数	0	1	2	3	4	$\geqslant 5$
频　　数	92	68	28	11	1	0

问，在单位时间内汽车过路辆数的分布可否看成普阿松分布？（$\alpha=0.10$）.

学 习 指 导

（一）本章概要

本章的内容可以概括成以下三个方面：

1　**基本概念**　本章介绍了数理统计学中最重要的一些基本概念，例如总体与子样；子样均值与子样方差；无偏估计、有效估计、一致估计；统计量及其分布等。

2　**理论结果**　首先要提到的是定理 5.1.2．它证明了，取自正态总体的样本的子样均值与子样方差是相互独立的，并且nS^2/σ^2服从$\chi^2(n-1)$．这一定理和它的几个推论奠定了显著性检验与区间估计理论的基础．这一定理的证明方法也是很典型的，值得细细玩味．

除此，还应提到的，就是子样均值$\overline{X}$作为总体均值μ的点估计的优良性质．

3　**具体方法**　在这一章里，对未知参数作点估计讲了矩法估计、最大似然估计两种方法．应该首先从直观上搞清楚它们的想法，尔后通过练习学会，并掌握这两种方法．

显著性检验与区间估计的内容有着密切的联系．教材中平行地讨论了种种不同的情况．应该注意到处理这类问题的程序是完全一样的，我们可以采用解剖麻雀的办法，例如弄懂U检验和σ已知时对未知的μ作区间估计的想法与作法，从中抽出一般的规律，在此基础上去掌握别的检验法就不困难了．

（二）内容说明

1　关于总体与子样的概念

理解总体的概念不能停留在它是某种对象的全体上，而应该把

它看成是具有确定的概率分布的随机变量．例如，研究吉粳60号水稻的株高，我们并不是把所有吉粳60号的水稻植株都拿来放在一起称为总体，而是只注意它的株高及株高落在某个区间的概率．如果株高的概率分布给定了，我们就说这一总体给定了．

对于初学者来说，接受"子样是随机向量"这一概念是比较困难的．他们一谈起取样就认为是取得了一组确定的数值．例如从一大批直径为 $8mm$ 的螺钉中抽取样本容量为1的样本，量其直径得 $7.9mm$．这确是一个具体的定数．但是，我们应该看到，在取样时，这一大批螺钉中的每一个螺钉均可能被取到，因而其直径的数值也会有所不同．从这样一个角度看来，一次抽样一个螺钉，量其直径应该是一个随机变量．不仅它所可能取的数值与总体 X 的可能值是一样的，而且它取值落在某个区间的概率与 X 取值于相同的区间的概率也是完全一样的．对于样本容量大于1的情形可作类似的分析．由此，我们应该把子样 $(X_1, X_2, \cdots, X_n)$ 看成一个 n 维随机向量，一次具体抽取所得到的数值 $(x_1, x_2, \cdots, x_n)$ 只是它的一个观测值．

建立"子样是随机向量"这一观念是至关重要的．在数理统计学的全部理论中，没有它可以说寸步难行．

为了强调子样与子样的观测值的区别，对于作为随机变量的子样，我们用大写拉丁字母 $(X_1, X_2, \cdots, X_n)$ 记之，它的观测值用小写拉丁字母 $(x_1, x_2, \cdots, x_n)$ 记之．对于子样的函数——统计量也有类似的记法，即用大写拉丁字母（或按照统计文献上习惯的写法，用某个固定的希腊字母）表示作为随机变量的统计量，对于统计量所取的值则用小写拉丁字母表示．例如 $\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$，$\overline{x}=\frac{1}{n}\sum_{i=1}^{n}x_i$，后者是前者的值．在不引起误会的情况下，有时同一个符号既可代表随机变量，又可代表它取的某个观测值．

2 关于定理5.1.2的证明

定理 5.1.2 要证明在正态总体下，子样均值$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$与子样方差$S^2=\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$是相互独立的。由于$S^2$的表达式中含有$\overline{X}$，因而直观上看不出这一结论。

在正态分布的情形下，相互独立与不相关是一回事。因此，我们只要能够证明，在适当的线性变换$Y=AX$下，有$DY=\sigma^2I_n$，而$\overline{X}$是Y_1的函数，S^2是Y_2，…，Y_n的函数即可。

如何去计算$DY=D(AX)$，这在定理 5.1.1 中已经讲了。这样一来，证明独立性的问题就化为计算方差矩阵的问题。用不多的代算演算就导出了结果。

显而易见，如果没有正态性的假定，没有相互独立与不相关是等价的结果，要想证明两个随机变量是相互独立的，显然会困难得多。

3 关于两种求点估计的方法

说得通俗一点，对未知参数进行点估计，就是依据子样所提供的信息去“猜”未知参数值的一种方法。

矩法估计的想法是很直观的。例如，我们想要知道吉粳60号水稻的平均株高，把所有这种品种的水稻植株的株高加以平均即可求得。但这是不可能的。自然会想到用一部分植株的株高平均作为它的估计值。一般地，我们可以用子样的各阶矩去估计总体的各阶矩。显而易见，这种办法是直观而又简便的。特别地，在对总体的均值与方差作矩法估计时，并不需要知道总体X的分布函数，而只要X的二阶矩存在即可。这种方法的缺点在于子样矩的表达式与总体X的分布函数$F(x;\theta)$的表达式无关，因而矩法没有充分利用$F(x;\theta)$对参数θ所提供的信息。

作为矩法估计的改进，费歇引进了最大似然估计。这种估计方法利用了总体X的分布函数$F(x;\theta)$的表达式及子样所提供的关于

θ 的信息。到目前为止，在理论上仍是参数点估计的一种最重要的方法。

最大似然法的想法也是很直观的。我们举一个例子。一个箱子里装有黑球、白球共 9 个，我们从中随机地、无放回地抽取 3 个球，发现恰有 2 个黑球。请猜一下箱子里有几个黑球，几个白球？

首先可以排除的可能是箱中恰有 2 个黑球的情况。箱中恰有 2 个黑球，你任取 3 个球就把这两个黑球全取出来了。这种事情实际上是很难迂见的。

为了给我们的猜测方法提供一个依据，下面作一点简单的计算。

当箱中恰有 i 个黑球时，取出的三个球中恰有 2 个黑球的事件记为 A_i，$i=2,3,4,5,6,7,8$，则

$$P(A_2)=\frac{\binom{2}{2}\binom{7}{1}}{\binom{9}{3}}=\frac{7}{84}=0.0833$$

$$P(A_3)=\frac{\binom{3}{2}\binom{6}{1}}{\binom{9}{3}}=\frac{18}{84}=0.2143$$

$$P(A_4)=\frac{\binom{4}{2}\binom{5}{1}}{\binom{9}{3}}=\frac{30}{84}=0.3571$$

$$P(A_5)=\frac{\binom{5}{2}\binom{4}{1}}{\binom{9}{3}}=\frac{40}{84}=0.4762$$

$$P(A_6)=\frac{\binom{6}{2}\binom{3}{1}}{\binom{9}{3}}=\frac{45}{84}=0.5357$$

$$P(A_7)=\frac{\binom{7}{2}\binom{2}{1}}{\binom{9}{3}}=\frac{42}{84}=0.5000$$

$$P(A_8)=\frac{\binom{8}{2}\binom{1}{1}}{\binom{9}{3}}=\frac{28}{84}=0.3333$$

哪一个事件的概率大，则在一次试验时发生的可能性也大。按照这样一个直观的想法，我们应当猜这时是事件 A_6 发生了，即箱中恰有 6 个黑球。最大似然法就是这样一个直观想法的一般化，我们选择 $\hat{\theta}$ 作为参数 θ 的估计，使得子样 $(x_1, x_2, \cdots, x_n)$ 发生的可能

性最大。

在连续型变量的情况，由于在一点取值的概率为0，因而我们应考虑（x_1，x_2，…，x_n）的一个邻域，即选取使得$\prod_{i=1}^{n} f(x_i;\theta)dx_i$达到最大的数值$\hat{\theta}$作为参数$\theta$的估计值。

4 关于公式$E\left(\sum_{i=1}^{n}(X_i-\overline{X})^2\right)=(n-1)\sigma^2$

我们知道，X的方差定义为$DX=E(X-\mu)^2$，其中$\mu=EX$。记$DX=\sigma^2$，有

$$E(X-\mu)^2=\sigma^2$$

如果X_1，X_2，…，X_n是从X中抽取的大小为n的子样，则对所有的$i=1$，2，…，n，均有

$$E(X_i-\mu)^2=\sigma^2$$

由此，

$$E\left[\sum_{i=1}^{n}(X_i-\mu)^2\right]=\sum_{i=1}^{n}E(X_i-\mu)^2=n\sigma^2$$

这表明平方和$\sum_{i=1}^{n}(X_i-\mu)^2$反映了数据$X_i$的波动的程度，显然它与求和的项数有关。为消除项数的影响，用n除之，有

$$E\left(\frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2\right)=\sigma^2$$

当总体均值μ未知的情形，自然地，我们会想到用$\overline{X}$去估计μ。这时$(X_i-\overline{X})^2$反映了数据X_i与$\overline{X}$偏离的程度。从而，$\sum_{i=1}^{n}(X_i-\overline{X})^2$，反映了一批数据$X_1$，$X_2$，…，$X_n$总的波动的大小。这个平方和也是$n$个加项的求和，为了消除项数的影响，初学者总认为应该用n去除，使得平均说来$\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$和$\sigma^2$相差不多。但是，这是办不到的，在正文中已经算出

$$E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2\right]=\frac{n-1}{n}\sigma^2$$

即平均说来，$\frac{1}{n}\sum_{i=1}^{n}(X_i-\overline{X})^2$ 总比σ^2偏小。为了得到σ^2的无偏估计，我们应该用$n-1$去除平方和$\sum_{i=1}^{n}(X_i-\overline{X})^2$，这样就有

$$E\left[\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\overline{X})^2\right]=\sigma^2$$

除因子 $n-1$ 习惯上称为平方和$\sum_{i=1}^{n}(X_i-\overline{X})^2$ 的自由度。它表明平方和$\sum_{i=1}^{n}(X_i-\overline{X})^2$中，由于存在一个独立的约束关系

$$\sum_{i=1}^{n}(X_i-\overline{X})=0$$

因而可以自由（独立）变动的变量个数为 $n-1$ 个。

作为公式 $E\left(\sum_{j=1}^{k}(X_j-\overline{X})^2\right)=(k-1)\sigma^2$ 的应用，我们有下述结果。

若总体X_1，X_2，…，X_m相互独立，有相同的方差σ^2，且X_{11}，X_{12}，…，X_{1k}；X_{21}，X_{22}，…，X_{2k}；…；X_{m1}，X_{m2}，…X_{mk}是分别从总体X_1，X_2，…，X_m中取出的子样，记

$$\overline{X}_{i\cdot}=\frac{1}{k}\sum_{j=1}^{k}X_{ij}\qquad \overline{X}=\frac{1}{m}\sum_{i=1}^{m}\overline{X}_i$$

则

(a) $E\left[\sum_{i=1}^{m}\sum_{j=1}^{k}(X_{ij}-\overline{X}_{i\cdot})^2\right]=m(k-1)\sigma^2$；

(b) 若进一步假定$EX_i=\mu+\alpha_i$，$i=1,2,\cdots,m$，且$\sum_{i=1}^{m}\alpha_i=0$，有

$$E\left[k\sum_{i=1}^{m}(X_{i\cdot}-\overline{X})^2\right]=(m-1)\sigma^2+k\sum_{i=1}^{m}\alpha_i^2$$

证明

$$(a)\ E\left[\sum_{i=1}^{m}\sum_{j=1}^{k}(X_{ij}-\overline{X}_{i\cdot})^2\right]=\sum_{i=1}^{m}E\left[\sum_{j=1}^{k}(X_{ij}-\overline{X}_{i\cdot})^2\right]$$

注意到，对任一给定的 i 均有

$$E\left[\sum_{j=1}^{k}(X_{ij}-\overline{X}_{i\cdot})^2\right]=(k-1)\sigma^2$$

所以

$$E\left[\sum_{i=1}^{m}\sum_{j=1}^{k}(X_{ij}-\overline{X}_{i\cdot})^2\right]=m(k-1)\sigma^2$$

(b) 首先，我们看到，对任一给定的 i 有

$$D\overline{X}_{i\cdot}=D\left(\frac{1}{k}\sum_{j=1}^{k}X_{ij}\right)=\frac{1}{k^2}\sum_{j=1}^{k}DX_{ij}=\frac{1}{k}\sigma^2$$

令 $Y_i=\overline{X}_{i\cdot}-\alpha_i$，则 $EY_i=E(\overline{X}_{i\cdot}-\alpha_i)$

$$=E\left[\frac{1}{k}\sum_{j=1}^{k}X_{ij}-\alpha_i\right]=\frac{1}{k}\sum_{j=1}^{k}EX_{ij}-\alpha_i=\mu+\alpha_i-\alpha_i=\mu,$$

$$DY_i=D(\overline{X}_{i\cdot}-\alpha_i)=D\overline{X}_i=\frac{\sigma^2}{k}$$

又 $\overline{Y}\triangleq\frac{1}{m}\sum_{i=1}^{m}Y_i=\frac{1}{m}\sum_{i=1}^{m}(X_i-\alpha_i)=\frac{1}{m}\sum_{i=1}^{m}\overline{X}_{i\cdot}-\frac{1}{m}\sum_{i=1}^{m}\alpha_i$

$=\overline{X}$，所以

$$E\left[\sum_{i=1}^{m}(Y_i-\overline{Y})^2\right]=(m-1)\cdot\frac{\sigma^2}{k}$$

但是，

$$E\left[k\sum_{i=1}^{m}(\overline{X}_{i\cdot}-\overline{X})^2\right]=E\left[k\sum_{i=1}^{m}(Y_i+\alpha_i-\overline{Y})^2\right]$$

$$=kE\left[\sum_{i=1}^{m}(Y_i-\overline{Y})^2\right]+k\sum_{i=1}^{m}\alpha_i^2+2k\sum_{i=1}^{m}\alpha_iE(Y_i-\overline{Y})$$

注意到$E\overline{Y}=E\left(\frac{1}{m}\sum_{i=1}^{m}Y_i\right)=\frac{1}{m}\sum_{i=1}^{m}EY_i=\mu$，有 $E(Y_i-\overline{Y})=EY_i-E\overline{Y}=\mu-\mu=0$，所以

$$E\left[k\sum_{i=1}^{m}(\overline{X}_{i\cdot}-\overline{X})^2\right]=k\cdot\frac{(m-1)\sigma^2}{k}+k\sum_{i=1}^{m}\alpha_i^2$$

$$=(m-1)\sigma^2+k\sum_{i=1}^{m}\alpha_i^2$$

5 关于显著性检验的想法

显著性检验的想法是很直观的。举个例子来讲，某人甲告诉某人乙，四川人都是高个子，身高在175cm左右（或使用概率论的语言讲，四川人的身高服从正态分布$N(175, 5^2)$。乙信不信呢？一次，乙正好有个机会乘飞机去四川。下飞机后乙迁见的第一个四川人身高只有158cm，这时，他就对甲的说法产生了怀疑。如果以后他碰上的25个四川人的平均身高不到170cm，则他干脆就不会相信甲的介绍了。乙是怎么考虑这一问题的呢？我们来分析一下。首先乙根据甲的介绍，对四川人的身高X提出了一个假设H_0: $X\sim N(175, 5^2)$。如果假设H_0真，则X取值在$[\mu-3\sigma, \mu+3\sigma]=[160, 190]$之外的概率将小于0.003。这是一个小概率事件，在实际生活中是罕见的。粗略地讲，观察1000次，这种事情发生不到3次。现在他作一次观察，这种事情就发生了，自然他会怀疑假设H_0的真确性。以后他观察了25人，平均身高不到170cm，这时他更有理由怀疑H_0的真确性了。因为，当H_0真时，$\overline{X}=\frac{1}{25}\sum_{i=1}^{25}X_i\sim N\left(175,\frac{5^2}{25}\right)$，由此$\overline{X}$取值在[170, 180]之外的概率

$$p=1-P(170\leqslant\overline{X}\leqslant180)$$

$$=1-P\left(\frac{170-175}{1}\leqslant\frac{\overline{X}-175}{1}\leqslant\frac{180-175}{1}\right)$$

$$=1-(\Phi(5)-\Phi(-5))=2(1-\Phi(5))\doteq0.0000006$$

这个数值更小。它表明作一千万组观察（每组25人），平均身高在〔170，180〕之外发生不到六次。现在，只进行一组观察，这样的小概率事件就发生了，自然，乙不会相信甲的介绍，从而拒绝假设 H_0。

这里乙进行推断的主要依据是“小概率事件在一次试验中实际上是不可能发生的”．人们在实际生活中处理问题是恪守这一原则的。例如，天上掉下一块陨石正好打在一头牛身上的概率是很小的，所以人们也就把这样的事情看成是实际上的不可能事件，不管在哪里都不会为此采取防范措施。究竟多小的概率才算小概率？这要看我们所要面临的问题而定。小概率事件并不是不可能事件，一旦它发生了，我们就要遭受损失。对于某些可能遭受大的损失的事情，我们不妨把小概率的数值定得特别小。对于一般的事情，人们总是取 $\alpha=0.05$ 或 $\alpha=0.01$ 作为小概率事件的概率。此外，确定小概率的数值还和提出假设时的慎重态度有关。如果我们的假设是经过深思熟虑提出来的，接受实践的检验时，就希望有更充足的理由才能对它表示怀疑。这时 α 就可以定得很小，并且还可多作几次观测才能最后作出决断。

显著性检验问题有一般的处理程序：

(a)提出假设 H_0；

(b)把子样空间分成不相交的两个部分 W 与 $\overline{W}$，使得当 H_0 真时，$P(X\in W)=\alpha$（α 是一个小概率）；

(c)当样本点（子样的观测值）落入 W 内时拒绝 H_0。

6　关于区间估计

区间估计问题是要用一个随机区间（这区间的端点均是随机变量）去套未知参数的问题。对于给定的置信度 $1-\alpha$，有

$$P(\hat{\theta}_1<\theta<\hat{\theta}_2)=1-\alpha$$

在教学中，我们发现有的学员出现这样的错误。例如，求出 θ 的置信区间为(13.1，13.3)后，他便写为

$$P(13.1<\theta<13.3)=0.95,$$

认为，未知参数θ落在(13.1，13.3)内的概率为95%。这种写法与这种看法都是错误的。产生这种错误的根本原因在于错误地把参数θ当成了随机变量。事实上，θ虽然未知，但却是一个常数，并非随机变量。比较常数的大小关系，只要学过比较实数的大小的同志就早已解决了。

那么$P(\hat{\theta}_1<\theta<\hat{\theta}_2)=1-\alpha$这一式子直观上是怎么一回事呢？不妨我们讲得具体一点。当σ^2已知时，对于正态总体均值μ的置信区间为

$$\left(\overline{X}-u_\alpha\frac{\sigma}{\sqrt{n}},\ \overline{X}+u_\alpha\frac{\sigma}{\sqrt{n}}\right)$$

即

$$P\left(\overline{X}-u_\alpha\frac{\sigma}{\sqrt{n}}<\mu<\overline{X}+u_\alpha\frac{\sigma}{\sqrt{n}}\right)=1-\alpha$$

这里，$\overline{X}-u_\alpha\frac{\sigma}{\sqrt{n}}$，$\overline{X}+u_\alpha\frac{\sigma}{\sqrt{n}}$均是子样的函数，从而它们都是随机变量。对于它们所取的可能值，有的能套住μ，有的不能套住μ。取$\alpha=0.05$时，能套住μ的情况所占的比例大约为95%，而套不住μ的情况所占的比例大约为5%。在套不住μ的这5%的情况，我们作出的断言是错的。这时犯了错误。如果我们用这种办法去作100次区间估计，出错的可能性大约是每100次中有5次。

如果读者对上述解释仍旧感到不具体的话，最好作一次验证性的练习。

例如，某厂维尼纶纤度服从正态$N(1.40,\ 0.048^2)$分布。现在设若纤度的均值$\mu=1.40$未知，用区间估计的办法去估计它。

取样，子样大小为9。得下述数据

1.36　1.49　1.43　1.41　1.37　1.40　1.32　1.42　1.47

算出子样均值为

$$\overline{x}=\frac{1}{9}(1.36+1.49+\cdots+1.47)=1.408$$

由此，置信度为0.95的置信区间为

$$\left(\overline{X}-1.96\times\frac{0.048}{3},\ \overline{X}+1.96\times\frac{0.048}{3}\right)$$

将$\overline{x}=1.408$代入得(1.377, 1.439)。我们看到这一区间套住了$\mu=1.40$，即断言 $1.377<\mu<1.439$ 是对的。

如果你抽得的子样，为另外 9 个数据：

1.42 1.40 1.41 1.37 1.46 1.36 1.37 1.27 1.37

由此也可以算出 $1.350<\mu<1.412$。这一说法也是对的。

当你进行 100 次抽样（每次抽得 9 个数据）时，相应地就得到关于 μ 的 100 个不等式，这其中大约有 5 个是不对的。例如，我们具体作了100次，其中有以下4个是不对的。

$1.316<\mu<1.378$

$1.330<\mu<1.392$

$1.409<\mu<1.471$

$1.416<\mu<1.478$

这 4 个式子均不正确。但出错的可能性的确很小，其余96个式子所作出的 μ 的范围都是对的。

7 关于非参数检验

关于分布函数的拟合检验仅是非参数检验一个方面的内容。此外，还有随机变量之间的独立性与相关性的检验。

用皮尔逊 χ^2 检验法作分布函数的拟合检验时，我们使用了统计量

$$\eta=\sum_{i=1}^{m}\frac{(\nu_i-np_i)^2}{np_i}$$

当假设 H_0 真时，频率$\frac{\nu_i}{n}$将稳定在概率 p_i 附近，从而 ν_i-np_i 将以概率 1 取值于 0 左右。这样一来，η 的值越小，反映拟合得越好。η 的值越大，反映拟合的情况越不好。因而这个统计量有可能作为检验假设 H_0：$F(x)=F_0(x)$ 的检验统计量。要真能用上，还必须求出 η 的分布。如果精确分布求不出，对固定的m，当$n\to\infty$

时，η的极限分布已知也行。这已由皮尔逊定理解决了。

8　关于统计问题的计算

在数理统计学中，一般说来，各种统计问题的计算量均较大。这在手算的情况下，极大地影响了统计方法的普及、推广与应用。但是，在电子计算器大规模普及的今天，这一问题已经不存在了。一般的电子计算器上均有统计运算的专用键钮，计算几十个数据的均值与标准差一、两分钟就足够了。因而本书对于一些简易的、直观的统计方法未作介绍。

习 题 解 答

1　（1）子样(X_1, X_2, X_3)的观测值(x_1, x_2, x_3)是一个三维（行）向量，其中$x_i=0$或1，$i=1, 2, 3$，所以，子样空间由下列$2^3=8$个点组成：(0,0,0)，(0,0,1)，(0,1,0)，(0,1,1)，(1,0,0)，(1,0,1)，(1,1,0)，(1,1,1)。

由简单随机子样的定义知，X_1, X_2, X_3是相互独立的随机变量，并且与总体X具有相同的概率分布，所以(X_1, X_2, X_3)的概率分布为

$$P[(X_1, X_2, X_3)=(x_1, x_2, x_3)]=p^k q^{3-k}$$

其中k为观测值(x_1, x_2, x_3)中1的个数。这是一种简略的记法，它代表了八个不同的等式。例如，当$(x_1, x_2, x_3)=(0, 1, 1)$时，其中1的个数$k=2$，有

$$P[(X_1, X_2, X_3)=(0, 1, 1)]=p^2q$$

它的计算过程是

$$P[(X_1, X_2, X_3)=(0, 1, 1)]=P[X_1=0, X_2=1, X_3=1]$$
$$=P((X_1=0)\cdot(X_2=1)\cdot(X_3=1))$$
$$=P(X_1=0)\cdot P(X_2=1)\cdot P(X_3=1)$$ （由X_1, X_2, X_3相互独立）

$$=q\cdot p\cdot p$$ （由X_1, X_2, X_3与X同分布）

$= q \cdot p^2$

（2）已知(0，1，1）是（X_1，X_3，X_2）的一个观测值，这时子样均值为

$$\overline{x} = \frac{1}{3}(0+1+1) = \frac{2}{3}$$

子样方差为

$$S^2 = \frac{1}{3}\sum_{i=1}^{3}(x_i - \overline{x})^2 = \frac{1}{3}\left(\sum_{i=1}^{3} x_i^2 - 3(\overline{x})^2\right)$$

$$= \frac{1}{3}\left(0^2 + 1^2 + 1^2 - 3\left(\frac{2}{3}\right)^2\right)$$

$$= \frac{2}{9}$$

2　（1）子样(X_1，X_2，…，X_5)的观测值(x_1，x_2，…，x_5)是一个五维（行）向量，而 $x_i \in (-\infty, \infty)$，$i = 1, 2, \cdots, 5$，所以，子样空间是所有可能的点($x_1$，$x_2$，…，$x_5$）的全体，即五维欧氏空间。

(X_1，X_2，…，X_5)的分布密度是

$$p(x_1, x_2, x_3, x_4, x_5) = \prod_{i=1}^{5} p_i(x_i)$$

$$= \frac{1}{(\sqrt{2\pi}\sigma)^5} e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{5}(x_i-\mu)^2}$$

（2）已知（0，1，－1，2，－2）是（X_1，X_2，…，X_5)的一个观测值，这时子样均值为

$$\overline{x} = \frac{1}{5}[0+1+(-1)+2+(-2)] = 0,$$

子样方差为

$$S^2 = \frac{1}{5}\sum_{i=1}^{5}(x_i - \overline{x})^2 = \frac{1}{5}\left(\sum_{i=1}^{5} x_i^2 - 5\overline{x}^2\right)$$

$$= \frac{1}{5}(0^2 + 1^2 + (-1)^2 + 2^2 + (-2)^2 - 5\times 0^2)$$

$$= 2$$

3　（1）证明

$$\overline{X}' = \frac{1}{n}\sum_{i=1}^{n} X_i' = \frac{1}{n}\sum_{i=1}^{n} (X_i + a) = \frac{1}{n}\sum_{i=1}^{n} X_i + \frac{1}{n}(na)$$

$$= \overline{X} + a$$

$$S'^2 = \frac{1}{n}\sum_{i=1}^{n} (X_i' - \overline{X}')^2 = \frac{1}{n}\sum_{i=1}^{n} [(X_i + a) - (\overline{X} + a)]^2$$

$$= \frac{1}{n}\sum_{i=1}^{n} (X_i - \overline{X})^2 = S^2$$

（2）将所有数据都减去35，得新数据

x_i'：-2，1，-1，1，1，0，-4，0，-2，-8

$$\overline{x} = \overline{x}' - a = \overline{x}' + 35 = -1.4 + 35 = 33.6$$

$$S^2 = S'^2 = \frac{1}{10}\sum_{i=1}^{10} (x_i' - \overline{x}')^2$$

$$= \frac{1}{10}[(-2+1.4)^2 + (1+1.4)^2 + \cdots + (-8+1.4)^2]$$

$$= \frac{1}{10} \times 72.4 = 7.24$$

4　已知X服从二项分布 $b(n, p)$，所以

$$\mu = np,\ \sigma^2 = npq = np(1-p)$$

由矩法得方程组

$$\begin{cases} \hat{n}\hat{p} = \hat{\mu} = \overline{X} \\ \hat{n}\hat{p}(1-\hat{p}) = \hat{\sigma}^2 = S^2. \end{cases}$$

解之，于是参数n与p的矩法估计分别为

$$\hat{n} = \frac{\overline{X}^2}{\overline{X} - S^2},\qquad \hat{p} = \frac{\overline{X} - S^2}{\overline{X}}$$

5　已知X服从参数为λ的普阿松分布，从而

$$\mu = EX = \lambda$$

由矩法得

$$\hat{\lambda} = \hat{\mu} = \overline{X}$$

所以，$\overline{X}$是未知参数 λ 的矩法估计。

又由于似然函数为

$$L(\lambda) = L(x_1, x_2, \cdots, x_n; \lambda) = \frac{\lambda^{\sum_{i=1}^{n} x_i}}{\prod_{i=1}^{n} x_i!} e^{-n\lambda}$$

$$\ln L(\lambda) = \left(\sum_{i=1}^{n} x_i\right)\ln\lambda - \sum_{i=1}^{n}\ln x_i! - n\lambda$$

$$\frac{\partial \ln L(\lambda)}{\partial \lambda} = \left(\sum_{i=1}^{n} x_i\right)\frac{1}{\lambda} - n$$

解$\frac{\partial \ln L(\lambda)}{\partial \lambda} = 0$，得

$$\hat{\lambda} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

所以，$\overline{X}$为 λ 的最大似然估计。

6　解（1）似然函数为

$$L(x_1, x_2, \cdots, x_n; \theta_1, \theta_2) = \prod_{i=1}^{n} p(x_i; \theta_1, \theta_2)$$

$$= \begin{cases} \dfrac{1}{(\theta_2 - \theta_1)^n} & \theta_1 \leqslant x_i (i = 1, 2, \cdots n) \leqslant \theta_2 \\ 0 & \text{其它} \end{cases}$$

显然 $\theta_2 - \theta_1$ 愈小，则 $L(\theta_1, \theta_2)$ 愈大，但 $\theta_2 - \theta_1$ 又不能小于 $\max\limits_{1\leqslant i\leqslant n} X_i - \min\limits_{1\leqslant i\leqslant n} X_i$。

所以得

$$\hat{\theta}_1 = \min_{1\leqslant i\leqslant n} X_i \qquad \hat{\theta}_2 = \max_{1\leqslant i\leqslant n} X_i。$$

（2）

$$L(\theta) = \frac{1}{\theta^n} \cdot e^{-\frac{\sum_{i=1}^{n} X_i}{\theta}}$$

$$\ln L(\theta) = -n\ln\theta - \frac{\sum_{i=1}^{n} X_i}{\theta}$$

由
$$\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{-n}{\theta} + \frac{\sum_{i=1}^{n} X_i}{\theta^2} = 0$$

得 $-n\theta + \sum_{i=1}^{n} X_i = 0$ $\therefore \hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$

7 解 似然函数为

$$L(p) = p^{x_1+x_2+\cdots+x_n}(1-p)^{n-(x_1+x_2+\cdots+x_n)}$$

$$= p^{n\overline{x}}(1-p)^{n-n\overline{x}}$$

$$\ln L(p) = n\overline{X}\ln p + (n-n\overline{X})\ln(1-p)$$

由

$$\frac{\partial \ln L(p)}{\partial p} = n\overline{X}\frac{1}{p} - (n-n\overline{X})\frac{1}{1-p} = 0$$

得 $(1-p)n\overline{X} = p(n-n\overline{X})$ $\therefore \hat{p} = \overline{X}$ (即频率)

8 解（1）似然函数为

$$L(\theta) = \begin{cases} 4(\theta - X_1)(\theta - X_2)\theta^{-4} & (0 < X_1, X_2 < \theta) \\ 0 & \text{其它} \end{cases}$$

$$\ln L(\theta) = \ln 4 + \ln(\theta - X_1) + \ln(\theta - X_2) - 4\ln\theta$$

由

$$\frac{\partial \ln L(\theta)}{\partial \theta} = \frac{1}{\theta - X_1} + \frac{1}{\theta - X_2} - \frac{4}{\theta} = 0$$

得

$$\theta(\theta - X_2) + \theta(\theta - X_1) - 4(\theta - X_1)(\theta - X_2) = 0$$

$$2\theta^2 - 3(X_1 + X_2)\theta + 4X_1X_2 = 0$$

所以

$$\hat{\theta} = \frac{3(X_1 + X_2) + \sqrt{9X_1^2 - 14X_1X_2 + 9X_2^2}}{4}$$

（2）由于

$$\mu = \int_0^\theta x \frac{2(\theta - x)}{\theta^2} dx = \frac{\theta}{3}$$

所以 μ 的最大似然估计为

$$\hat{\mu} = \frac{\hat{\theta}}{3}$$

9　证明

$$EX = \int_0^\theta xp(x;\ \theta)dx = \int_0^\theta \frac{x}{\theta}dx = \frac{\theta}{2}$$

$$E(\hat{\theta}) = E(2\overline{X}) = E\left(\frac{2}{n}\sum_{i=1}^{n} X_i\right) = \frac{2}{n}\sum_{i=1}^{n} EX_i = \frac{2}{n}\cdot n\cdot\frac{\theta}{2}$$

$$= \theta$$

10　解　因为

$$E(\hat{\mu}) = E(c_1X_1 + c_2X_2 + \cdots + c_nX_n) = (c_1 + c_2 + \cdots + c_n)\mu = \mu$$

$$D(\hat{\mu}) = D(c_1X_1 + c_2X_2 + \cdots + c_nX_n) = (c_1^2 + c_2^2 + \cdots + c_n^2)\sigma^2$$

所以 c_i 应满足 $c_1 + c_2 + \cdots + c_n = 1$ 且使 $c_1^2 + c_2^2 + \cdots + c_n^2$ 达到最小。将 $c_1 = 1 - c_2 - c_3 - \cdots - c_n$ 代入 $c_1^2 + c_2^2 + \cdots + c_n^2$ 中，令各偏导数等于 0 可解得

$$c_1 = c_2 = c_3 = \cdots = c_n = \frac{1}{n}$$

11　证明　因为 $D\overline{X} = E\overline{X}^2 - (E\overline{X})^2$，所以

$$E(\overline{X}^2) = D\overline{X} + (E\overline{X})^2 = \frac{\sigma^2}{n} + \mu^2 > \mu^2$$

12　证明

$$E(S_0^2) = \frac{1}{n}E\left[\sum_{i=1}^{n}(X_i - \mu)^2\right] = \frac{1}{n}\sum_{i=1}^{n}E(X_i - \mu)^2$$

$$= \frac{1}{n}\cdot n\cdot\sigma^2 = \sigma^2$$

13　证明

$$E(\hat{\mu}) = \mu \qquad \text{当} \lim_{n\to+\infty}\sum_{i=1}^{n}c_i^2 \longrightarrow 0 \text{时}$$

$$D(\widehat{\mu}) = \sigma^2 \sum_{i=1}^{n} c_i^2 \to 0$$

所以，根据切贝晓夫不等式，有

$$P(|\widehat{\mu} - \mu| \geqslant \varepsilon) \leqslant \frac{D(\widehat{\mu})}{\varepsilon^2} \to 0$$

14　解　设 k_1，k_2 为所求，则

$$E(k_1 \widehat{\theta}_1 + k_2 \widehat{\theta}_2) = (k_1 + k_2)\theta = \theta$$

$$\therefore \quad k_1 + k_2 = 1 \tag{1}$$

又　$D(k_1 \widehat{\theta}_1 + k_2 \widehat{\theta}_2) = (2k_1^2 + k_2^2) D\widehat{\theta}_2$

所以应使 $2k_1^2 + k_2^2$ 达到最小．将 $k_1 = 1 - k_2$ 代入 $2k_1^2 + k_2^2$ 中，令其导数等于 0 可解得

$$k_1 = \frac{1}{3}, \qquad k_2 = \frac{2}{3}.$$

15　（1）证明　设 $F(x)$ 为 X 的分布函数，则

$$F(x) = \begin{cases} 1 & x \geqslant \theta \\ \dfrac{x}{\theta} & 0 \leqslant x < \theta \\ 0 & x < 0 \end{cases}$$

令　$Y = \max\limits_{1 \leqslant i \leqslant 3} X_i$，　　$Z = \min\limits_{1 \leqslant i \leqslant 3} X_i$

则 Y 的分布函数与分布密度分别为

$$F_y(x) = [F(x)]^3$$

$$f_y(x, \theta) = \begin{cases} 3\left(\dfrac{x}{\theta}\right)^2 \cdot \dfrac{1}{\theta} & 0 \leqslant x < \theta \\ 0 & \text{其它} \end{cases}$$

Z 的分布函数与分布密度分别为

$$F_Z(x) = 1 - [1 - F(x)]^3$$

$$f_z(x, \theta) = \begin{cases} 3\left(1 - \dfrac{x}{\theta}\right)^2 \dfrac{1}{\theta} & 0 \leqslant x < \theta \\ 0 & \text{其它} \end{cases}$$

所以

$$EY=\frac{3}{\theta^3}\int_0^\theta x^3dx=\frac{3}{4}\theta$$

$$E\left(\frac{4}{3}\max_{1\leq i\leq 3}X_i\right)=\theta$$

$$EZ=\frac{3}{\theta^3}\int_0^\theta x(\theta-x)^2dx$$

$$=\frac{3}{\theta^3}\int_0^\theta(\theta^2x-2\theta x^2+x^3)dx=\frac{1}{4}\theta$$

$$E(4\min_{1\leq i\leq 3}X_i)=\theta$$

所以$\frac{4}{3}\max\limits_{1\leq i\leq 3}X_i$与$4\min\limits_{1\leq i\leq 3}X_i$都是$\theta$的无偏估计。

(2) $DY=EY^2-(EY)^2=\frac{3}{\theta}\int_0^\theta x^2\left(\frac{x}{\theta}\right)^2dx-\left(\frac{3}{4}\theta\right)^2$

$$=\frac{3}{\theta^3}\int_0^\theta x^4dx-\frac{9}{16}\theta^2=\frac{3}{5}\theta^2-\frac{9}{16}\theta^2=\frac{3}{80}\theta^2$$

$$\therefore\quad D\left(\frac{4}{3}\max_{1\leq i\leq 3}X_i\right)=\frac{16}{9}DY=\frac{1}{15}\theta^2$$

$$DZ=EZ^2-(EZ)^2=\frac{3}{\theta}\int_0^\theta x^2\left(1-\frac{x}{\theta}\right)^2dx-\left(\frac{1}{4}\theta\right)^2$$

$$=\frac{3}{\theta^3}\int_0^\theta x^2(\theta-x)^2dx-\frac{1}{16}\theta^2=\frac{1}{10}\theta^2-\frac{1}{16}\theta^2=\frac{3}{80}\theta^2$$

$$\therefore\quad D(4\min_{1\leq i\leq 3}X_i)=16DZ=16\cdot\frac{3}{80}\theta^2=\frac{3}{5}\theta^2$$

$$\therefore\quad D\left(\frac{4}{3}Y\right)<D(4Z)$$

$\therefore\quad \frac{4}{3}\max\limits_{1\leq i\leq 3}X_i$ 比 $4\min\limits_{1\leq i\leq 3}X_i$ 有效。

16. 解

$$E[(\hat{\theta}_1-\theta)^2]=D(\hat{\theta}_1)=D(\overline{X})=\frac{1}{n}$$

$$E[(\hat{\theta}_2-\theta)^2]=(0.05-\theta)^2$$

由于 $(0.05-\theta)^2 \leqslant (0.05)^2 = \frac{1}{400}$.

因此，当 $n<400$ 时

$$E[(\hat{\theta}_1-\theta)^2] = \frac{1}{n} > \frac{1}{400} > [E(\hat{\theta}_2-\theta)^2]$$

当 $n>400$ 时，两者大小关系不确定.

17　解　这里假设 $H_0: \mu=12(g)$，由于 $\sigma=1$ 已知，所以用 U 统计量进行检验. 计算 U 的观测值

$$|u| = \left|\frac{\overline{x}-\mu_0}{\sigma}\sqrt{n}\right| = \left|\frac{12.5-12}{1}\cdot 10\right| = 5$$

因为 $\alpha=0.10$，所以 $u_\alpha=1.65$. 由于 $|u|=5>1.65$，所以否定假设 H_0. 即认为设备更新前后产品的平均重量有变化.

18　解　由于 σ 未知，所以用 t 统计量进行检验. 计算 t 的观测值

$$|t| = \left|\frac{\overline{x}-\mu_0}{s}\sqrt{n-1}\right| = \left|\frac{42-50}{5}\sqrt{19}\right|$$
$$= 6.97$$

这里 $\alpha=0.01$，自由度为 19，查表得 $t_\alpha=2.861$，由于 $|t| = 6.97>t_\alpha$，所以否定假设 H_0.

19　解　这里假设 $H_0: \mu_1=\mu_2$，由于 $\sigma_1=\sigma_2$ 未知，所以用 t 统计量进行检验. 计算 t 的观测值

$$|t| = \left|\frac{\overline{x}_A-\overline{x}_B}{\sqrt{\frac{n(s_1^2+s_2^2)2n}{n^2(2n-2)}}}\right| = \left|\frac{2.33-0.75}{\sqrt{\frac{10(3.61+2.89)20}{100(20-2)}}}\right|$$
$$= \frac{1.58}{0.85} \doteq 1.86$$

这里 $\alpha=0.01$，自由度为18，查表得 $t_\alpha=2.878$，由于 $|t|=1.86<2.878$，所以不能否定假设 H_0，可以认为 A，B 对延长睡眠时间的平均效果没有显著差异.

20　解　这里假设 $H_0: \mu=7.20$，由于 σ 未知，所以用 t 统计量进行检验.

为此先求$\overline{x}$，s^*。

$\overline{x}=7.9 \quad s^{*2}=\frac{1}{8}(0.8^2+0.6^2+0.2^2+0.1^2+0.1^2+0.2^2+0.6^2+$

$+1.1^2+0.3^2)$

$\doteq 0.345$

$s^* \doteq 0.587$

计算 t 的观测值

$$|t|=\left|\frac{\overline{x}-\mu_0}{s^*}\sqrt{n}\right|=\left|\frac{7.9-7.2}{0.587}\sqrt{9}\right|$$

$$\doteq 3.58$$

这里 $\alpha=0.05$，自由度为 8，查表得 $t_\alpha=2.306$

由于$|t|=3.58>2.306$。所以否定假设H_0，即可以认为该患者红血球平均直径与健康人有显著差异.

21 解 这里假设 $H_0:\mu_1=\mu_2$，由于 $\sigma_1=\sigma_2$ 未知，所以用 t 统计量进行检验。

$$t=\frac{\overline{X}-\overline{Y}}{\sqrt{\frac{(mS_1^2+nS_2^2)(m+n)}{mn(m+n-2)}}}$$

先计算$\overline{x}$，$\overline{y}$，ms_1^2，ns_2^2 得

$\overline{x}=0.27 \quad \overline{y}=0.13$

$ms_1^2=10s_1^2=0.253 \qquad m=10$

$ns_2^2=0.064 \qquad n=11$

代入得 t 值

$$|t|=\left|\frac{0.27-0.13}{\sqrt{\frac{(0.253+0.064)21}{11\times 10\times 19}}}\right|\doteq 2.5$$

这里$\alpha=0.05$,自由度为19,查表得 $t_\alpha=2.093$，由于$|t|=2.5>2.093$,所以否定假设 H_0，即可认为处理前后含脂率的总体均值有显著变化。

22 解 要求 $P(|\overline{X}-\mu_0|\geqslant k\sigma)\leqslant 0.05$，等价于要求

$$P\left(\frac{|\overline{X}-\mu_0|}{\sigma}\sqrt{n}\geqslant\sqrt{n}k\right)\leqslant 0.05$$

由于 $\frac{\overline{X}-\mu_0}{\sigma}\sqrt{n}$ 服从 $N(0,1)$，所以 $\sqrt{n}k\geqslant 1.96$，

解得 $n\geqslant\frac{(1.96)^2}{k^2}\doteq\frac{3.84}{k^2}$

当 $k=1$时，$n\geqslant 4$

当 $k=\frac{1}{2}$时，$n\geqslant\frac{3.84}{\frac{1}{4}}=15.36$，即 $n\geqslant 16$

当 $k=\frac{1}{3}$时 $n\geqslant\frac{3.84}{\frac{1}{9}}=34.56$，即$n\geqslant 35$

23 解 由于 μ 未知，所以用统计量

$$\chi^2=\frac{\sum_{i=1}^{n}(X_i-\overline{X})^2}{\sigma^2}$$

进行检验．计算 χ^2 的观测值

$$\chi^2=\frac{652.8}{64}=10.2$$

这里 $\alpha=0.05$，自由度为 7，查表得

$$\chi^2_{\frac{\alpha}{2}}=16.01$$

$$\chi^2_{1-\frac{\alpha}{2}}=1.69$$

由于 $1.69<10.2<16.01$

所以不能否定假设 H_0．

24 解 由于 $\mu=10$已知，所以用统计量

$$\chi^2=\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}$$

进行检验．计算 χ^2 的观测值

$$\chi^2=(9.4-10)^2+(11.3-10)^2+(8.7-10)^2$$
$$+(10.6-10)^2+(9.7-10)^2=4.19$$

这里 $\alpha=0.05$，自由度为 5 ，查表得

$$\chi^2_{\frac{\alpha}{2}}=12.83$$

$$\chi^2_{1-\frac{\alpha}{2}}=0.83$$

由于$0.83<4.19<12.83$，所以不能否定假设 H_0。

25　解　用统计量

$$F=\frac{S_1^{*2}}{S_2^{*2}}$$

进行检验．当 H_0 成立时，F 服从自由度为(8，10)的 F 分布．计算 F 的观测值

$$F=\frac{96/8}{45/10}\doteq 2.67$$

这里 $\alpha=0.02$，自由度为(8，10)，查 F 表得

$$F_{\frac{\alpha}{2}}=F_{0.01}=5.06$$

由于$F=2.67<5.06$，所以不能否定假设 H_0。

26　解　当 σ 已知时，μ 的置信区间应为

$$\left(\overline{X}-u_\alpha\frac{\sigma}{\sqrt{n}},\quad \overline{X}+u_\alpha\frac{\sigma}{\sqrt{n}}\right)$$

（1）这里 $\alpha=0.05$．所以 $u_\alpha=1.96$，置信区间为

$$\left(12-1.96\frac{4}{12},\quad 12+1.96\frac{4}{12}\right)$$

即　　(11.347，　12.653)

（2）$\alpha=0.01$　　　$u_\alpha=2.58$，置信区间为

$$\left(12-2.58\frac{4}{12},\quad 12+2.58\frac{4}{12}\right)$$

即　　(11.14，　　12.86)

27　解　当 σ 未知时，μ 的置信区间应为

$$\left(\overline{X}-t_\alpha\frac{S}{\sqrt{n-1}},\quad \overline{X}+t_\alpha\frac{S}{\sqrt{n-1}}\right)$$

或写成

$$\left(\overline{X}-t_\alpha\frac{S^*}{\sqrt{n}},\quad \overline{X}+t_\alpha\frac{S^*}{\sqrt{n}}\right)$$

（1）$\alpha=0.05$，自由度为 4，$t_\alpha=2.776$

所以置信区间为

$$\left(12.3-2.776\frac{0.2}{2},\quad 12.3+2.776\frac{0.2}{2}\right)$$

即 $(12.3-0.278,\quad 12.3+0.278)$

（2）$\alpha=0.05$，自由度为20，$t_\alpha=2.086$

所以置信区间为

$$\left(13.2-2.086\sqrt{\frac{5}{20}},\quad 13.2+2.086\sqrt{\frac{5}{20}}\right)$$

即 $(13.2-1.043,\quad 13.2+1.043)$

（3）$\alpha=0.05$，自由度为35，$t_\alpha=2.03$

所以置信区间为

$$\left(15.2-2.03\sqrt{\frac{68.4}{36}},\quad 15.2+2.03\sqrt{\frac{68.4}{36}}\right)$$

即 $(15.2-2.798,\quad 15.2+2.798)$

28 解 当μ未知时，σ^2的置信区间应为

$$\left(\frac{nS^2}{\chi_2^2},\ \frac{nS^2}{\chi_1^2}\right)$$

其中 $\chi_1^2=\chi^2_{1-\frac{\alpha}{2}}$ $\chi_2^2=\chi^2_{\frac{\alpha}{2}}$

（1）$\alpha=0.10$ $\frac{\alpha}{2}=0.05$ $\chi^2_{1-\frac{\alpha}{2}}=13.85$

$\chi^2_{\frac{\alpha}{2}}=36.415$ $S^2=\frac{1}{n}\sum_{i=1}^{n}X_i^2-\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)^2=0.156$

$nS^2=3.9$，所以置信区间为

$$\left(\frac{3.90}{36.42},\ \frac{3.90}{13.85}\right)\doteq(0.11,\ 0.28)$$

（2）$ns^2=14.916$，$\alpha=0.02$，$\frac{\alpha}{2}=0.01$

$\chi^2_{1-\frac{\alpha}{2}}=3.05$ $\quad\chi^2_{\frac{\alpha}{2}}=24.73$ 所以置信区间为

$$\left(\frac{14.916}{24.73},\ \frac{14.916}{3.05}\right)\doteq(0.60,\ 4.89)$$

29 解 $\alpha=0.05$ $\quad u_\alpha=1.96$ $\quad$ 置信区间为

$$\left(\overline{X}-1.96\frac{10}{\sqrt{n}},\qquad \overline{X}+1.96\frac{10}{\sqrt{n}}\right)$$

若其长为5，则 $2\times1.96\frac{10}{\sqrt{n}}=5$ $\quad n\doteq61$

若 $\alpha=0.01$ $\quad u_\alpha=2.58$ $\quad$ 置信区间为

$$\left(\overline{X}-2.58\frac{10}{\sqrt{n}},\ \overline{X}+2.58\frac{10}{\sqrt{n}}\right)$$

若其长为5，则 $2\times2.58\frac{10}{\sqrt{n}}=5$ $\quad n\doteq106$

30 解 $\overline{x}=\frac{1}{10}\sum_{i=1}^{10}x_i=52$

$$s^2=\frac{1}{10}\sum_{i=1}^{10}x_i^2-(52)^2\doteq96.2$$

$s\doteq9.80$

$\alpha=0.10$ $\quad\frac{\alpha}{2}=0.05$ $\quad\chi^2_{1-\frac{\alpha}{2}}=3.33$ $\quad\chi^2_{\frac{\alpha}{2}}=16.92$

σ 的置信区间应为

$$\left(\frac{\sqrt{n}\,S}{\chi_{\frac{\alpha}{2}}}\ \frac{\sqrt{n}\,S}{\chi_{1-\frac{\alpha}{2}}}\right)$$

将计算数值代进去得置信区间为

$$\left(\frac{\sqrt{10}\cdot9.8}{\sqrt{16.92}},\qquad \frac{\sqrt{10}\cdot9.8}{\sqrt{3.33}}\right)=(7.54,\ 17.02)$$

31 解 由于方差 σ_1^2、σ_2^2 已知，所以置信区间应为

$$\left(\overline{X}-\overline{Y}-u_\alpha\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}},\qquad \overline{X}-\overline{Y}+u_\alpha\sqrt{\frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}}\right)$$

这里$\sigma_1^2=50$，$\sigma_2^2=60$，$\overline{x}=80$，$\overline{y}=70$，$m=10$，$n=30$，$\alpha=0.05$，$u_\alpha=1.96$，代入则得置信区间为

$$\left(80-70-1.96\sqrt{\frac{50}{10}+\frac{60}{30}},\quad 80-70+1.96\sqrt{\frac{50}{10}+\frac{60}{30}}\right)$$

$$=(4.81,\ 15.19)$$

32　解　置信区间应近似为

$$\left(\overline{X}-u_\alpha\sqrt{\overline{X}(1-\overline{X})/n},\quad \overline{X}+u_\alpha\sqrt{\overline{X}(1-\overline{X})/n}\right)$$

其中　$\overline{x}=\frac{64}{100}$　$u_\alpha=1.96$　$n=100$ 代入得置信区间为

$$(0.64-1.96\sqrt{0.64(1-0.64)/100},$$

$$0.64+1.96\sqrt{0.64(1-0.64)/100})$$

$$=(0.55,\quad 0.73)$$

33　解　由于$\overline{X}=\frac{1}{n_1}\sum_{i=1}^{n_1}X_i$，　$\overline{Y}=\frac{1}{n_2}\sum_{i=1}^{n_2}Y_i$　分别服从正态分布 $N\left(\mu_1,\ \frac{\sigma_1^2}{n_1}\right)$，$N\left(\mu_2,\ \frac{\sigma_2^2}{n_2}\right)$，所以$\overline{X}-\overline{Y}$服从正态分布 $N\left(\mu_1-\mu_2,\ \frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_1}\right)$，因此

$$U=\frac{\overline{X}-\overline{Y}-(\mu_1-u_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}$$

服从 $N(0,1)$。对于置信度0.90，查表可求满足下式的 u_α

$$0.90=P\left(-u_\alpha<\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}}<u_\alpha\right)$$

解括弧内的不等式得

$$0.90=P\left(\overline{X}-\overline{Y}-1.65\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}<\mu_1-\mu_2<\overline{X}-\overline{Y}\right.$$

$$\left.+1.65\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$

因此置信区间为

$$\left(\overline{X}-\overline{Y}-1.65\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}},\quad \overline{X}-\overline{Y}+1.65\sqrt{\frac{\sigma_1^2}{n_1}+\frac{\sigma_2^2}{n_2}}\right)$$

34 首先计算$\overline{x}$与s^2，由题给的数据，利用电子计算器立即算出

$\overline{x}=343.64 \doteq 343.5$，

$s \doteq 4$。

假设H_0：$F(x)=N(343.5,\ 4^2)$

把数据分成七组，主要计算结果列成下表

区　间	$(-\infty,333.5)$	$(333.5,337.5)$	$(337.5,341.5)$	$(341.5,345.5)$
ν_i	1	4	23	42
p_i	0.0062	0.0606	0.2417	0.383
np_i	0.62	6.06	24.17	38.3
$(\nu_i-np_i)^2$	0.1444	4.2436	1.3689	13.69
$\frac{(\nu_i-np_i)^2}{np_i}$	0.2329	0.7003	0.0566	0.3574

区　间	$(345.5,\ 349.5)$	$(349.5,\ 353.5)$	$(353.5,\ +\infty)$
ν_i	21	8	1
p_i	0.2417	0.0606	0.0062
np_i	24.17	6.06	0.62
$(\nu_i-np_i)^2$	10.0489	3.7636	0.1444
$\frac{(\nu_i-np_i)^2}{np_i}$	0.4158	0.6211	0.2329

计算χ^2值得

$$\chi^2=\sum_{i=1}^{7}\frac{(\nu_i-np_i)^2}{np_i}=2.6170$$

查 χ^2 分布表，自由度为 4，$\alpha=0.10$，$\chi^2_{0.10}=7.779$，现在$2.617<7.779$，故不能拒绝 H_0，即不否认服从正态分布的假设。

35 以X记一直到掷出白色的一面与地面接触时的投掷次数，则X服从几何分布。假定正四面体是均匀的，则出现白色一面触地的概率为$\frac{1}{4}$，从而

$$P(X=k)=\left(\frac{3}{4}\right)^{k-1}\cdot\frac{1}{4},\ k=1,\ 2,\ 3,\ \cdots$$

由此，

$$p_1=P(X=1)=0.25,\quad p_2=P(X=2)=\frac{3}{16}\doteq0.1875,$$

$$p_3=P(X=3)=\frac{9}{64}\doteq0.1406,\ p_4=P(X=4)=\frac{27}{256}\doteq0.1055,$$

$$p_5=P(X\geqslant5)=\frac{81}{256}\doteq0.3164$$

用皮尔逊 χ^2 检验，将主要计算结果列成下表

次　数	1	2	3	4	$\geqslant5$
v_i	56	48	32	28	36
p_i	0.25	0.1875	0.1406	0.1055	0.3164
np_i	50	37.5	28.12	21.10	63.28
$(v_i-np_i)^2$	36	110.25	15.0544	47.61	744.1984
$\frac{(v_i-np_i)^2}{np_i}$	0.72	2.94	0.5354	2.2564	11.7604

于是，

$$\chi^2=\sum_{i=1}^{5}\frac{(v_i-np_i)^2}{np_i}=18.2122$$

查 χ^2 分布表，自由度为 $5-1=4$，$\alpha=0.05$，$\chi^2_{0.05}=9.488$，现在$18.2122>9.488$，故拒绝 H_0，即认为此四面体是不均匀的。

36 首先算出

$$\overline{x}=\frac{1}{200}(0\times 92+1\times 68+2\times 28+3\times 11+4\times 1)=0.805$$

$$\doteq 0.8$$

由此可假设H_0：X服从参数为0.8的普阿松分布。

当H_0成立时，查普阿松分布表得

$p_0=P(X=0)=0.44933$，$p_1=P(X=1)=0.35946$

$p_2=P(X=2)=0.14379$，$p_3=P(X=3)=0.03834$

$p_4=P(X=4)=0.00766$，$p_5=P(X\geqslant 5)=0.001411$

用皮尔逊χ^2检验。将主要计算结果列成下表

辆　数	0	1	2	3	4	≥5
ν_i	92	68	28	11	1	0
p_i	0.1493	0.3595	0.1438	0.0383	0.0077	0.0014
np_i	89.86	71.90	28.76	7.66	1.54	0.28
$(\nu_i-np_i)^2$	4.5796	15.21	0.5776	11.1556	0.2916	0.0784
$\frac{(\nu_i-np_i)^2}{np_i}$	0.05096	0.2115	0.0201	1.4563	0.1893	0.28

于是

$$\chi^2=\sum_{i=0}^{5}\frac{(\nu_i-np_i)^2}{np_i}=2.2082$$

查 χ^2 分布表，自由度为 $6-2=4$，$\alpha=0.10$，$\chi^2_{0.10}=7.779$，现在$2.208<7.779$，故不能拒绝 H_0，即不能否认关于X服从参数为0.8的普阿松分布的假设。

第六章　方差分析与回归分析

在生产实践与科学实验中，往往需要处理试验数据．例如，进行多次试验，得到的试验结果各不相同．我们要问，试验结果间的差异是什么原因造成的？具体地说，是由于试验条件的改变引起的呢？还是由于试验误差的影响？在数理统计学中，方差分析回答这一类问题．又如，考察人的年龄与血压之间的关系．由于同一年龄的人的血压不尽相同，因而不可能存在类似于圆的面积$S=\pi r^2$那样的确定性关系，但它们之间仍然存在某种相关的关系．回归分析就是研究相关关系的一种数学工具．

在这一章里，我们仅介绍常用的一些最基本的方差分析与回归分析的方法．

§6.1　单因子方差分析

今后，我们把生产实践与科学实验中的结果，如产品的性能、成本、产量等统称为指标，影响指标的因素称为因子，用A、B、C、…表示．因子在试验中所取的不同状态称为水平，因子A的不同水平用A_1，A_2，…表示．例如，一项作物栽培试验，要考虑不同的施氮肥量（30斤、50斤、70斤）对产量的影响．这时施氮肥量就是因子，而30斤、50斤、70斤就是它的三个不同的水平．又如，一项化工试验，要考虑不同的温度（低温、常温、高温）和不同的压力（1个大气压、5个大气压、20个大气压、50个大气压）对产品产量的影响．这时温度因子取三个水平，压力因子取四个水平．

在实践中，最简单的试验只考虑一个因子对试验结果的影响，

称为单因子试验．单因子方差分析就是讨论单因子试验的统计分析方法．对于多因子试验的情况，分析的方法与两因子试验的方差分析是类似的，只不过是更加繁杂而已．因此，我们将在§ 6.2 着重讨论两因子方差分析．

（一）方法概说

在给出一元方差分析方法的完整的数学描述之前，我们结合一个实例概括地讲一下它的思路．

例如，一项作物栽培试验，考虑施氮肥量（A）的不同水平：A_1(30斤)，A_2(50斤)，A_3(70斤)，A_4(90斤) 对作物产量的影响．重复三次，得数据如下：

数据 重复 \ 水平	A_1	A_2	A_3	A_4
1	750	790	770	810
2	780	765	830	830
3	810	815	800	790
平均	780	790	800	810

我们要问施氮肥量这一因子对作物产量是否有显著的影响？

在两种极端的情况下，这一问题直观上是很容易回答的．一种情况是，各个不同的施肥水平下作物的平均产量相差很大（比方说，200斤，400斤，1000斤），这时，很明显地，施氮肥量对产量的影响显著；另一种情况，在各个不同的施肥水平下的平均产量相差甚微（比方说，600斤，599斤，601斤，601.5斤），这时，谁都知道，多施氮肥对提高产量不起作用了．

大量的情况介于上述两种情况之间，直观上并不容易回答．看

我们给出的数据，四种不同施肥水平下的平均产量有差异，大体上施肥越多，产量越高。但是我们看到，由于随机误差的存在，既使在同一施肥水平下不同小区的产量数据波动也不小，这样自然会对上述看法表示怀疑，平均产量间的差异也许是由于随机误差造成的吧！

解决这一问题可以用显著性检验的方法。记在 A_i 水平下的理论产量为 μ_i，$i=1,2,3,4$。可以提出假设，在各施肥水平下的平均产量间没有显著差异，即假设 H_0：$\mu_1=\mu_2=\mu_3=\mu_4$。如果我们能够导出一个可以用来检验这一假设的统计量 F，那么这一问题就解决了。按照显著性检验的一般程序，对给定的 α，可找出一个F_α，使得$P(F\geqslant F_\alpha)=\alpha$。如果根据样本的观测值算出的 F 值大于F_α，我们就拒绝H_0，认为平均产量间有差异，否则就不拒绝H_0。

为了导出统计量 F，我们先具体分析一下。

记在A_i水平下的产量为Y_i，$i=1,2,3,4$，可以假定 $Y_i\sim N(\mu_i,\sigma^2)$，其中$\mu_i$，$\sigma^2$ 均未知。这时，在 A_1 水平下，三次重复所得到的数据（750，780，810）就是从 Y_1总体取出的样本容量为 3 的子样 (Y_{11},Y_{12},Y_{13}) 的一组观测值，由此，我们可以用子样均值

$$\overline{Y}_{1\cdot}=\frac{1}{3}(Y_{11}+Y_{12}+Y_{13})$$

和子样方差

$$S^{*2}_1=\frac{1}{3-1}\sum_{j=1}^{3}(Y_{1j}-\overline{Y}_{1\cdot})^2$$

分别去估计μ_1和σ^2。

类似地，可以分别用$\overline{Y}_{2\cdot}$，$\overline{Y}_{3\cdot}$，$\overline{Y}_{4\cdot}$去估计μ_2，μ_3，μ_4，这里

$$\overline{Y}_{i\cdot}=\frac{1}{3}(Y_{i1}+Y_{i2}+Y_{i3})\quad i=1,2,3,4$$

还可以用

$$S^{*2}_2=\frac{1}{3-1}\sum_{j=1}^{3}(Y_{2j}-\overline{Y}_{2\cdot})^2$$

$$S^{*2}_3=\frac{1}{3-1}\sum_{j=1}^{3}(Y_{3j}-\overline{Y}_{3\cdot})^2$$

$$S_4^{*2}=\frac{1}{3-1}\sum_{j=1}^{3}(Y_{4j}-\overline{Y}_{4\cdot})^2 \text{去估计} \sigma^2.$$

由于诸 S_i^{*2} 均是σ^2的无偏估计，所以

$$E\left[\frac{1}{4}\sum_{i=1}^{4}S_i^{*2}\right]=E\left[\frac{1}{4}\cdot\frac{1}{3-1}\sum_{i=1}^{4}\sum_{j=1}^{3}(Y_{ij}-\overline{Y}_{i\cdot})^2\right]$$
$$=\sigma^2$$

令 $S_e^2=\sum_{i=1}^{4}\sum_{j=1}^{3}(Y_{ij}-\overline{Y}_{i\cdot})^2$ 这个平方和反映了随机误差的大小，称为误差平方和。从而

$$\frac{1}{4(3-1)}S_e^2$$

是σ^2的无偏估计。

当$Y_i\sim N(\mu_i,\ \sigma^2)$时，$\overline{Y}_{i\cdot}=\frac{1}{3}\sum_{j=1}^{3}(Y_{ij})\sim N(\mu_i,\frac{\sigma^2}{3})$，假设$\mu_1=\mu_2=\mu_3=\mu_4=\mu$（这表明施氮肥量的不同对产量没有影响），则对一切$i$，有$\overline{Y}_{i\cdot}\sim N(\mu,\ \frac{\sigma^2}{3})$由于

$$\overline{Y}=\frac{1}{4}(\overline{Y}_{1\cdot}+\overline{Y}_{2\cdot}+\overline{Y}_{3\cdot}+\overline{Y}_{4\cdot})$$

所以

$$E\left[\frac{1}{4-1}\sum_{i=1}^{4}(\overline{Y}_{i\cdot}-\overline{Y})^2\right]=\frac{\sigma^2}{3}$$

$$E\left[\frac{3}{4-1}\sum_{i=1}^{4}(\overline{Y}_{i\cdot}-\overline{Y})^2\right]=\sigma^2$$

记 $S_A^2=3\sum_{i=1}^{4}(\overline{Y}_{i\cdot}-\overline{Y})^2$称为$A$的偏差平方和。它主要反映了$A$的不同水平下平均产量间的差异。当假设 $\mu_i=\mu$，$i=1,2,3,4$真时，$\frac{1}{4-1}S_A^2$也是σ^2的无偏估计。

这样一来，当假设H_0：$\mu_1=\mu_2=\mu_3=\mu_4$成立时，

$$\frac{1}{4-1}S_A^2$$

$$\frac{1}{4\ (3-1)}S_e^2$$

均是 σ^2的无偏估计，因而它们应该相差不多，或它们的商

$$F=\frac{S_A^2/(4-1)}{S_e^2/4(3-1)}$$

不能太大．如果由观测值算出的 F 值过大，则拒绝假设H_0，即认为不同的施氮肥量对产量有显著影响．

如何判定 F 的值是否过大，应该有一个标准（临界值）．为此必须求出在H_0成立的条件下，F 所服从的分布．

下面，我们从单因子方差分析的数学模型开始，给出完整的数学描述．

（二）数学模型

对因子 A 的 m 个水平各重复地做 k 次试验，数据如下：

表 6.1.1

数据 / 水平 / 重复序号	A_1	A_2	……	A_m
1	Y_{11}	Y_{21}	……	Y_{m1}
2	Y_{12}	Y_{22}	……	Y_{m2}
⋮	⋮	⋮	⋮	⋮
k	Y_{1k}	Y_{2k}	……	Y_{mk}

我们的任务是根据这组数据进行统计分析，主要回答下面一些问题．

1　对因子 A 各水平的作用如何估计；

2　各个水平的不同是否会引起试验结果的显著差异；

3　选取哪个水平对生产最有利等．

我们知道，在试验过程中由于试验误差（随机误差）的干扰，即使 k 次试验都是在同一水平下进行的，这 k 次试验的结果也会有所不同。我们把没有任何误差干扰的理想情况下，相应于因子 A 的第 i 个水平 A_i 的指标值记为 μ_i，并称它为在 A_i 下的工程平均。在 A_i 下第 j 次重复试验时所出现的试验误差记为 ε_{ij}。通常试验误差 ε_{ij} 具有下述性质：

1 ε_{ij} 是一个随机变量；

2 ε_{ij} 有时取正值，有时取负值，但都在 0 附近取值，取得正值的可能性与取得负值的可能性相同，并且与 0 出现大偏差的可能性很小；

3 对于所有的 i，j，ε_{ij} 的方差相等。

因此，我们可以假定 ε_{ij} 是相互独立的，具有相同方差 σ^2 的正态变量，且 $E(\varepsilon_{ij})=0$，$i=1, 2, \cdots, m$，$j=1, 2, \cdots, k$，这样一来，表（6.1.1）中的数据可表为

$$Y_{ij}=\mu_i+\varepsilon_{ij}, \quad i=1,2,\cdots, m; \ j=1,2,\cdots,k \qquad (6.1.1)$$

以后，我们称（6.1.1）为有 m 个水平的单因子试验的数学模型。

通常，为了便于分析各个水平所起的作用，把参数 μ_i 写成

$$\mu_i=\mu+\alpha_i \qquad (6.1.2)$$

其中 $\mu=\frac{1}{m}\sum_{i=1}^{m}\mu_i$ 称为一般平均，并且把

$$\alpha_i=\mu_i-\mu \qquad i=1,2,\cdots,m$$

称为第 i 个水平的效应，显然有

$$\sum_{i=1}^{m}\alpha_i=\sum_{i=1}^{m}(\mu_i-\mu)=0$$

把（6.1.2）式代入（6.1.1）式得

$$Y_{ij}=\mu+\alpha_i+\varepsilon_{ij}, i=1,2,\cdots,m; \ j=1,2,\cdots,k \quad (6.1.3)$$

这样一来，我们可以把统计分析的任务归结为：

1 对 μ，α_i 作参数估计，$1\leqslant i\leqslant m$；

2 判断因子 A 的各个水平效应有无显著差异，即检验统计假

设H_0：$\alpha_1=\alpha_2=\cdots=\alpha_m=0$是否成立；

3　对各水平下的工程平均作区间估计。

（三）参数估计

在这一段，我们要求出μ_i，μ，α_i的估计$\hat{\mu}_i$，$\hat{\mu}$，$\hat{\alpha}_i(i=1,2,\cdots,m)$，并研究它们所具有的性质。

由(6.1.1)知，Y_{ij}是随机变量，且$E(Y_{ij})=\mu_i$，$D(Y_{ij})=\sigma^2$。因此，对任一个固定的 i，Y_{i1}，Y_{i2}，……，Y_{ik} 是从正态总体$N(\mu_i,\ \sigma^2)$中取出的一个容量为k的子样。很自然地，我们会想到用子样均值

$$\overline{Y}_{i\cdot}=\frac{1}{k}\sum_{j=1}^{k}Y_{ij}\qquad i=1,2,\cdots,m$$

作为μ_i的估计$\hat{\mu}_i$。显然，$E(\hat{\mu}_i)=E(\overline{Y}_{i\cdot})=\mu_i$，即$\overline{Y}_{i\cdot}$是$\mu_i$的无偏估计。

由此出发，可推出如下的一些估计量。因为

$$\mu=\frac{1}{m}\sum_{i=1}^{m}\mu_i$$

令

$$\hat{\mu}=\frac{1}{m}\sum_{i=1}^{m}\hat{\mu}_i=\frac{1}{mk}\sum_{i=1}^{m}\sum_{j=1}^{k}Y_{ij}=\overline{Y}$$

为μ的估计量。由

$$E(\hat{\mu})=E\left(\frac{1}{m}\sum_{i=1}^{m}\hat{\mu}_i\right)=\frac{1}{m}\sum_{i=1}^{m}E(\hat{\mu}_i)$$

$$=\frac{1}{m}\sum_{i=1}^{m}\mu_i=\mu$$

知，$\hat{\mu}=\overline{Y}$是μ的无偏估计。令

$$\hat{\alpha}_i=\hat{\mu}_i-\hat{\mu}=\overline{Y}_{i\cdot}-\overline{Y}\qquad i=1,2,\cdots,m$$

作为α_i的估计量。由

$$E(\hat{\alpha}_i)=E(\hat{\mu}_i-\hat{\mu})=E(\hat{\mu}_i)-E(\hat{\mu})=\mu_i-\mu=\alpha_i$$

知$\hat{\alpha}_i$是α_i的无偏估计，且

$$\sum_{i=1}^{m}\hat{\alpha}_i=\sum_{i=1}^{m}(\overline{Y}_{i\cdot}-\overline{Y})=\sum_{i=1}^{m}\overline{Y}_{i\cdot}-m\overline{Y}$$

$$=\frac{1}{k}\sum_{i=1}^{m}\sum_{j=1}^{k}Y_{ij}-\frac{1}{k}\sum_{i=1}^{m}\sum_{j=1}^{k}Y_{ij}=0$$

注意到，

$$Y_{ij}-\overline{Y}=(Y_{ij}-\overline{Y}_{i\cdot})+(\overline{Y}_{i\cdot}-\overline{Y})$$

令

$$e_{ij}=Y_{ij}-\overline{Y}_{i\cdot},\qquad i=1,2,\cdots,m;\ j=1,2,\cdots,k$$

则

$$Y_{ij}=\overline{Y}+(\overline{Y}_{i\cdot}-\overline{Y})+e_{ij}$$

$$=\hat{\mu}+\hat{\alpha}_i+e_{ij}\qquad i=1,2,\cdots,m;\ j=1,2,\cdots,k$$

(6.1.4)

比较（6.1.4）与（6.1.3）式，可以看出它们具有类似的形式，$\hat{\mu}$，$\hat{\alpha}_i$恰是μ，α_i的无偏估计，而e_{ij}以后将会看到，也正好反映了误差ε_{ij}。

（四）显著性检验

表6.1.1的数据相当于从m个正态总体$N(\mu_i,\sigma^2)$分别抽出的子样。由于A_i对试验结果的影响由它的效应α_i所刻划，因而判断因于A的不同水平对于试验结果是否具有显著影响就归结为检验假设

$$H_0:\ \alpha_1=\alpha_2=\cdots=\alpha_m=0 \qquad (6.1.5)$$

是否成立了。

为了导出检验假设（6.1.5）的统计量，我们首先分析一下引起诸Y_{ij}波动的原因。当假设（6.1.5）成立时，诸Y_{ij}的波动纯粹是随机误差引起的；当假设（6.1.5）不成立时，除了随机误差外，各个水平效应的不同也引起Y_{ij}的波动。由此，这就启示我们设法把反映诸Y_{ij}波动程度的量加以分解，使得一个加项表示随机

误差的影响，另一个加项还反映各个水平效应的不同。这就是所谓的平方和分解的方法，而检验假设（6.1.5）的统计量亦可由此导出。

令

$$S_T^2=\sum_{i=1}^{m}\sum_{j=1}^{k}(Y_{ij}-\overline{Y})^2$$

其中，$\overline{Y}=\frac{1}{mk}\sum_{i=1}^{m}\sum_{j=1}^{k}Y_{ij}$是所有数据的总平均。称$S_T^2$为总的偏差平方和。它反映了数据$Y_{ij}$的波动程度的大小。注意到，

$$\begin{aligned}(Y_{ij}-\overline{Y})^2&=[(Y_{ij}-\overline{Y}_{i\cdot})+(\overline{Y}_{i\cdot}-\overline{Y})]^2\\&=(Y_{ij}-\overline{Y}_{i\cdot})^2+2(Y_{ij}-\overline{Y}_{i\cdot})(\overline{Y}_{i\cdot}-\overline{Y})\\&\quad+(\overline{Y}_{i\cdot}-\overline{Y})^2\end{aligned}$$

且

$$\begin{aligned}&\sum_{i=1}^{m}\sum_{j=1}^{k}(\overline{Y}_{i\cdot}-\overline{Y})(Y_{ij}-\overline{Y}_{i\cdot})\\&=\sum_{i=1}^{m}[(\overline{Y}_{i\cdot}-\overline{Y})\sum_{j=1}^{k}(Y_{ij}-\overline{Y}_{i\cdot})]\\&=\sum_{i=1}^{m}[(\overline{Y}_{i\cdot}-\overline{Y})(\sum_{j=1}^{k}Y_{ij}-k\overline{Y}_{i\cdot})]=0\end{aligned}$$

所以，

$$\begin{aligned}\sum_{i=1}^{m}\sum_{j=1}^{k}(Y_{ij}-\overline{Y})^2&=\sum_{i=1}^{m}\sum_{j=1}^{k}(Y_{ij}-\overline{Y}_{i\cdot})^2\\&\quad+k\sum_{i=1}^{m}(\overline{Y}_{i\cdot}-\overline{Y})^2\end{aligned}$$

记

$$S_e^2=\sum_{i=1}^{m}\sum_{j=1}^{k}(Y_{ij}-\overline{Y}_{i\cdot})^2$$

$$S_A^2=k\sum_{i=1}^{m}(\overline{Y}_{i\cdot}-\overline{Y})^2$$

则有

$$S_T^2=S_A^2+S_e^2 \tag{6.1.6}$$

称为平方和的分解公式。

为了弄清S_A^2，S_e^2的含义，我们计算S_A^2，S_e^2的均值。

$$E(S_e^2)=E\left(\sum_{i=1}^{m}\sum_{j=1}^{k}(Y_{ij}-\overline{Y}_{i\cdot})^2\right)$$
$$=\sum_{i=1}^{m}E\left(\sum_{j=1}^{k}(Y_{ij}-\overline{Y}_{i\cdot})^2\right)$$

注意到

$$\frac{1}{k}\sum_{j=1}^{k}(Y_{ij}-\overline{Y}_{i\cdot})^2$$

是从第i个正态总体$N(\mu_i,\sigma^2)$取出的容量为k的子样（Y_{i1}; Y_{i2}, …, Y_{ik}）的子样方差，由（5.2.5）式知，

$$E\left[\sum_{j=1}^{k}(Y_{ij}-\overline{Y}_{i\cdot})^2\right]=(k-1)\sigma^2,$$

所以

$$E(S_e^2)=\sum_{i=1}^{m}E\left(\sum_{j=1}^{k}(Y_{ij}-\overline{Y}_{i\cdot})^2\right)$$
$$=m(k-1)\sigma^2 \tag{6.1.7}$$

当$Y_{ij}\sim N(\mu_i,\sigma^2)$时，$\overline{Y}_{i\cdot}-\alpha_i\sim N\left(\mu,\frac{\sigma^2}{k}\right)$

注意到

$$\frac{1}{m}\sum_{i=1}^{m}(\overline{Y}_{i\cdot}-\alpha_i)=\frac{1}{m}\sum_{i=1}^{m}\overline{Y}_{i\cdot}=\overline{Y}$$

则

$$\frac{1}{m}\sum_{i=1}^{m}[(\overline{Y}_{i\cdot}-\alpha_i)-\overline{Y}]^2$$

是子样（$\overline{Y}_{1\cdot}-\alpha_1$，$\overline{Y}_{2\cdot}-\alpha_2$，…，$\overline{Y}_{m\cdot}-\alpha_m$）的子样方差。由（5.2.5）式知，

$$E\left(\sum_{i=1}^{m}(\overline{Y}_{i\cdot}-\alpha_i-\overline{Y})^2\right)=(m-1)\frac{\sigma^2}{k}$$

但是，

$$(\overline{Y}_{i\cdot}-\overline{Y})^2=(\overline{Y}_{i\cdot}-\alpha_i-\overline{Y}+\alpha_i)^2$$

$$= (\overline{Y}_{i\cdot} - a_i - \overline{Y})^2 + a_i^2 + 2a_i(\overline{Y}_{i\cdot} - a_i - \overline{Y})$$

并且

$$E(\overline{Y}_{i\cdot} - a_i - \overline{Y}) = E(\overline{Y}_{i\cdot} - a_i) - E\left(\frac{1}{m}\sum_{i=1}^{m}\overline{Y}_{i\cdot}\right)$$
$$= \mu - \mu = 0,$$

所以，

$$E(S_A^2) = kE\left(\sum_{i=1}^{m}(\overline{Y}_{i\cdot} - \overline{Y})^2\right)$$
$$= kE\left(\sum_{i=1}^{m}(\overline{Y}_{i\cdot} - a_i - \overline{Y})^2\right) + k\sum_{i=1}^{m}a_i^2$$
$$+ 2kE\left(\sum_{i=1}^{m}a_i(\overline{Y}_{i\cdot} - a_i - \overline{Y})\right)$$
$$= k(m-1)\cdot\frac{\sigma^2}{k} + k\sum_{i=1}^{m}a_i^2$$
$$+ 2k\sum_{i=1}^{m}a_i E(\overline{Y}_{i\cdot} - a_i - \overline{Y})$$
$$= (m-1)\sigma^2 + k\sum_{i=1}^{m}a_i^2 \qquad (6.1.8)$$

由（6.1.7）与（6.1.8）两式看出，平方和S_e^2反映了随机误差的影响，它的均值等于$m(k-1)\sigma^2$；而平方和S_A^2，当假设H_0真时，也反映了误差的影响，但当H_0不真时，还反映了A的不同水平效应间的差异。因此，比值

$$F = \frac{\dfrac{S_A^2}{m-1}}{\dfrac{S_e^2}{m(k-1)}} \qquad (6.1.9)$$

当假设H_0不成立时有偏大的趋势。在下一段我们将严格证明，在假设H_0真的条件下，统计量

$$F = \frac{S_A^2/(m-1)}{S_e^2/m(k-1)} \sim F(m-1,\ m(k-1)) \qquad (6.1.10)$$

对于假设 H_0：$a_1 = a_2 = \cdots = a_m = 0$，按照通常的程序进行显著

性检验。给定显著水平 α，查 F 分布表中，自由度为 $m-1, m(k-1)$ 时所相应的值 F_α，满足 $P(F \geqslant F_\alpha) = \alpha$。当由实测数据算出的 F 值 $\geqslant F_\alpha$ 时，拒绝 H_0，认为在显著水平 α 下，因子 A 的各水平效应间有显著差异。

*（五）柯赫伦（cochran）定理及其推论

在这一段，我们给出有名的柯赫伦定理。它在本章若干重要的理论结果的证明中将起重要的作用。如果读者代数准备不够，初读时可跳过这一段，只要会用它的结论即可。

定理6.1.1　设 A_i 是 n 阶实对称矩阵，记 $Q_i = x' A_i x$，$rkA_i = n_i$，$i = 1, 2, \cdots, k$，并且

$$x'x = \sum_{i=1}^{k} x' A_i x$$

则存在正交变换 $Z = Cx$，使得

$$x' A_i x = \sum_{j=n_1+\cdots+n_{i-1}+1}^{n_1+\cdots+n_{i-1}+n_i} Z_j^2,\quad i = 1, 2, \cdots, k \qquad (6.1.11)$$

成立的充要条件为

$$n = \sum_{i=1}^{k} n_i \qquad (6.1.12)$$

证明　必要性是显然的。事实上，若这样一个正交变换存在，则

$$x'x = \sum_{j=1}^{n_1+n_2+\cdots+n_k} Z_j^2$$

左端是秩为 n 的二次型，右端是秩为 $n_1+n_2+\cdots+n_k$ 的二次型，所以 $n = \sum_{i=1}^{k} n_i$。

反之，由于 A_i 为对称阵，则存在非奇异阵 P_i，使得

$$A_i = P_i \begin{pmatrix} D_i & O \\ O & O \end{pmatrix} P_i'$$

其中，D_i 为 n_i 阶对角矩阵，对角元素为 $+1$ 或为 -1。将 P_i 分块为 $(\underset{n\times n_i}{C_i} \vdots \underset{n\times(n-n_i)}{G})$，则

$$A_i = C_i D_i C_i' \qquad i = 1, 2, \cdots, k.$$

由此，

$$x'x = \sum_{i=1}^{k} x' A_i x = \sum_{i=1}^{k} x' C_i D_i C_i' x$$

$$= x'\left(\sum_{i=1}^{k} C_i D_i C_i'\right) x$$

$$= x'(C_1 C_2 \cdots C_k)\begin{pmatrix} D_1 & & O \\ & D_2 & \\ & & \ddots \\ O & & D_k \end{pmatrix}\begin{pmatrix} C_1' \\ C_2' \\ \vdots \\ C_k' \end{pmatrix} x$$

记 $C = (C_1 C_2 \cdots C_k)$， $D = \begin{pmatrix} D_1 & & O \\ & D_2 & \\ & & \ddots \\ O & & D_k \end{pmatrix}$，上式表明

$$I_n = CDC' \tag{6.1.13}$$

因为$rk(AB) \leqslant min(rkA,\ rkB)$，我们有

$$n = rkI_n = rkCDC' \leqslant rkC \leqslant \sum_{i=1}^{k} n_i = n \Rightarrow rkC = n$$

这表明矩阵 C 是 n 阶满秩方阵，故 C^{-1} 存在。用 C^{-1}，$(C')^{-1}$ 乘(6.1.13)式两端得

$$D = C^{-1} CDC'(C')^{-1} = C^{-1} I_n (C')^{-1} = C^{-1}(C^{-1})'$$

由正定阵的判定条件知D为正定阵，它的对角元素应全部大于 0，即D的对角元素不能为 -1，只能全部为 $+1$，即$D = I_n$。从而

$$CC' = I_n$$

即C是正交阵。

令

$$Z = \begin{pmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_k \end{pmatrix} = \begin{pmatrix} C_1' \\ C_2' \\ \vdots \\ C_k' \end{pmatrix} x$$

则得

$$x' A_i x \quad x' C_i D_i C_i' x = x' C_i C_i' x$$

$$=Z_i'Z_i=\sum_{j=n_1+\cdots+n_{i-1}+1}^{n_1+\cdots+n_{i-1}+n_i} Z_j^2$$

定理6.1.2（柯赫伦定理）设 n 维随机向量$Y\sim N(0,I_n)$，且

$$Y'Y=\sum_{i=1}^{k} Y'A_iY,\quad rkA_i=n_i,\quad i=1,2,\cdots,k,$$

则

$$Y'A_iY,\qquad i=1,2,\cdots,k$$

为相互独立的$\chi^2(n_i)$变量的充要条件是

$$n=\sum_{i=1}^{k} n_i.$$

证明　已知$Y\sim N(0,I_n)$，即Y_1，Y_2，…，Y_n相互独立，且服从相同的$N(0,1)$分布。$Y'Y=\sum_{i=1}^{n} Y_i^2$ 是诸 Y_i 变量的平方和，因此，$Y'Y\sim\chi^2(n)$。

正交变换$Z=CY$是Y的线性变换．因此Z仍然服从正态分布，且

$$E(Z)=E(CY)=CEY=0,$$

$$D(Z)=D(CY)=C\cdot DY\cdot C'=CI_nC'=I_n$$

这表明$Z\sim N(0,I_n)$，即$z_1,z_2,\cdots,z_n$ 相互独立，且服从相同的正态分布$N(0,1)$。

注意到 $\sum_{j=n_1+\cdots+n_{i-1}+1}^{n_1+n_2+\cdots+n_i} Z_j^2$ 是n_i个相互独立的正态 $N(0,1)$ 变量的平方和，所以具有$\chi^2(n_i)$分布。

这样一来

$$Y'A_iY,\qquad i=1,2,\cdots,k$$

是否为相互独立的 $\chi^2(n_i)$ 变量就归结为是否存在一个正交变换 $Z=CY$，使得

$$Y'A_iY=\sum_{j=n_1+\cdots+n_{i-1}+1}^{n_1+n_2+\cdots+n_i} Z_j^2\qquad i=1,2,\cdots,k$$

由定理6.1.1知，这一组等式成立的充要条件是

$$n=\sum_{i=1}^{k}n_i.$$

这就完成了本定理的证明。

现在，我们利用柯赫伦定理来讨论上一段留下的几个理论问题。

推论 在单因子试验的数学模型下，如果假设 H_0：$\alpha_1=\alpha_2=\cdots\alpha_m=0$ 成立，则

$$S_e^2/\sigma^2\sim\chi^2(mk-m)$$

$$S_A^2/\sigma^2\sim\chi^2(m-1)$$

且$\dfrac{S_e^2}{\sigma^2}$与$\dfrac{S_A^2}{\sigma^2}$相互独立，

$$\frac{S_A^2/(m-1)}{S_e^2/(mk-m)}\sim F(m-1,\ mk-m)$$

证明 在假设H_0：$\alpha_1=\alpha_2=\cdots=\alpha_m=0$成立时，有$\mu_1=\mu_2=\cdots=\mu_m=\mu$。由(6.1.6)得

$$\sum_{i=1}^{m}\sum_{j=1}^{k}(Y_{ij}-\mu)^2=\sum_{i=1}^{m}\sum_{j=1}^{k}[(Y_{ij}-\overline{Y})+(\overline{Y}-\mu)]^2$$

$$=\sum_{i=1}^{m}\sum_{j=1}^{k}(Y_{ij}-\overline{Y})^2+\sum_{i=1}^{m}\sum_{j=1}^{k}(\overline{Y}-\mu)^2$$

$$+2(\overline{Y}-\mu)\sum_{i=1}^{m}\sum_{j=1}^{k}(Y_{ij}-\overline{Y})$$

$$=\sum_{i=1}^{m}\sum_{j=1}^{k}(Y_{ij}-\overline{Y})^2+mk(\overline{Y}-\mu)^2$$

$$=S_A^2+S_e^2+mk(\overline{Y}-\mu)^2$$

对于 S_e^2，它有m个线性关系$\sum_{j=1}^{k}(Y_{ij}-\overline{Y}_{i\cdot})=0$，$i=1,2,\cdots,m$，所以$S_e^2$的秩为$mk-m$；

对于S_A^2它只包含一个线性关系$k\sum_{i=1}^{m}(\overline{Y}_{i\cdot}-\overline{Y})=0$，所以$S_A^2$的秩是$m-1$。

$mk(\overline{Y}-\mu)^2$的秩为1。

这三项都是非负定二次型，且$(mk-m)+(m-1)+1=mk$，所以由柯赫伦定理知，S_A^2/σ^2与S_e^2/σ^2分别为相互独立的$\chi^2(m-1)$，$\chi^2(mk-m)$变量，再由第三章定理3.5.11知，统计量

$$F=\frac{S_A^2/(m-1)}{S_e^2/(mk-m)}\sim F(m-1,\ mk-m)$$

（六）工程平均的估计

现在讨论在因子A的各水平下的工程平均应该如何估计．分两种情况．

1　显著性检验的结论是因子A不显著．这时假设H_0成立，即$\alpha_1=\alpha_2=\cdots\alpha_m=0$．所以由(6.1.3)式知

$$Y_{ij}=\mu+\varepsilon_{ij},\qquad i=1,2,\cdots,m;\ j=1,2,\cdots,k$$

这表明，A的各水平下的工程平均都是μ．表6.1.1中的数据Y_{ij}可以看成从正态总体$N(\mu,\sigma^2)$中取出的大小为mk的子样．对μ进行“预测”就变成对μ作点估计与区间估计的问题了．这在上一章已作了充分的讨论．

2　显著性检验的结论是因子A显著．这时，工程平均为

$$\mu_i=\mu+\alpha_i,\qquad i=1,2,\cdots,m.$$

它的点估计是$\hat{\mu}_i=\overline{Y}+(\overline{Y}_{i\cdot}-\overline{Y})=\overline{Y}_{i\cdot}$，　$i=1,2,\cdots,m$.

下面对μ_i作区间估计．由于$\overline{Y}_{i\cdot}=\frac{1}{k}(Y_{i1}+Y_{i2}+\cdots+Y_{ik})$是$k$个同正态$N(\mu_i,\sigma^2)$的相互独立随机变量的线性函数，因此，$\overline{Y}_{i\cdot}$服从正态分布$N\left(\mu_i,\frac{\sigma^2}{k}\right)$，从而$(\overline{Y}_{i\cdot}-\mu_i)/\frac{\sigma}{\sqrt{k}}\sim N(0,1)$．注意到

$$\frac{k(\overline{Y}_{i\cdot}-\mu_i)^2}{\sigma^2}$$

是正态 $N(0,1)$ 变量的平方，所以它服从 $\chi^2(1)$。我们已经证出 $S_e^2/\sigma^2 \sim \chi^2(mk-m)$，由柯赫伦定理可知 $k(\overline{Y}_{i\cdot}-\mu_i)^2/\sigma^2$ 与 S_e^2/σ^2 相互独立，因此，

$$\frac{k(\overline{Y}_{i\cdot}-\mu_i)^2}{S_e^2/(mk-m)} \sim F(1, mk-m)$$

这样，按照通常的区间估计的作法就可以给出 μ_i 的区间估计了

给定置信度 $1-\alpha$，从 F 分布表中查出 F_α 使得

$$P\left(\frac{k(\overline{Y}_{i\cdot}-\mu_i)^2}{S_e^2/(mk-m)} < F_\alpha\right) = 1-\alpha$$

$$P\left(\overline{Y}_{i\cdot} - \sqrt{\frac{S_e^2 F_\alpha}{mk^2-mk}} < \mu_i < \overline{Y}_{i\cdot} + \sqrt{\frac{S_e^2 F_\alpha}{mk^2-mk}}\right) = 1-\alpha$$

令 $\delta = \sqrt{\dfrac{S_e^2 F_\alpha}{mk^2-mk}}$ 上式表明，在置信度 $1-\alpha$ 下，μ_i 的置信区间是

$$(\overline{Y}_{i\cdot}-\delta, \overline{Y}_{i\cdot}+\delta)$$

（七）计算表格

我们看到方差分析的计算量是比较大的。为了避免出错，我们应该设计相应的计算表格，并给出一些能够减少计算误差的途径。

首先，我们把计算平方和的公式变形，这样可以较好地减少计算误差。

引理　设 x_1，x_2，…，x_m 是 m 个数，它们的算术平均数是

$$\overline{x} = \frac{1}{m}\sum_{i=1}^{m} x_i$$

则有

$$\sum_{i=1}^{m}(x_i-\overline{x})^2 = \sum_{i=1}^{m} x_i^2 - m\overline{x}^2 \qquad (6.1.11)$$

证明　由于$(x_i-\overline{x})^2=x_i^2-2\overline{x}x_i+\overline{x}^2$，两端对 i 求和，得

$$\sum_{i=1}^{m}(x_i-\overline{x})^2=\sum_{i=1}^{m}x_i^2-2\overline{x}\sum_{i=1}^{m}x_i+m\overline{x}^2$$

由于$\sum_{i=1}^{m}x_i=m\overline{x}$，等式右端中

$$-2\overline{x}\sum_{i=1}^{m}x_i+m\overline{x}^2=-2\overline{x}\cdot\overline{x}\cdot m+m\overline{x}^2=-m\overline{x}^2$$

所以

$$\sum_{i=1}^{m}(x_i-\overline{x})^2=\sum_{i=1}^{m}x_i^2-m\overline{x}^2$$

利用上述引理，注意到

$$\overline{Y}_{i\cdot}=\frac{1}{k}\sum_{j=1}^{k}Y_{ij}$$

$$\overline{Y}=\frac{1}{m}\sum_{i=1}^{m}\overline{Y}_{i\cdot}$$

所以

$$S_e^2=\sum_{i=1}^{m}\sum_{j=1}^{k}(Y_{ij}-\overline{Y}_{i\cdot})^2=\sum_{i=1}^{m}\left[\sum_{j=1}^{k}Y_{ij}^2-k\overline{Y}_{i\cdot}^2\right]$$

$$=\sum_{i=1}^{m}\left[\sum_{j=1}^{k}Y_{ij}^2-\frac{1}{k}\left(\sum_{j=1}^{k}Y_{ij}\right)^2\right]$$

$$=\sum_{i=1}^{m}\sum_{j=1}^{k}Y_{ij}^2-\frac{1}{k}\sum_{i=1}^{m}\left(\sum_{j=1}^{k}Y_{ij}\right)^2$$

$$S_A^2=k\sum_{i=1}^{m}\left(\overline{Y}_{i\cdot}-\overline{Y}\right)^2=k\left[\sum_{i=1}^{m}\overline{Y}_{i\cdot}^2-m\overline{Y}^2\right]$$

$$=k\left[\frac{1}{k^2}\sum_{i=1}^{m}\left(\sum_{j=1}^{k}Y_{ij}\right)^2-\frac{m}{(mk)^2}\left(\sum_{i=1}^{m}\sum_{j=1}^{k}y_{ij}\right)^2\right]$$

$$=\frac{1}{k}\sum_{i=1}^{m}\left(\sum_{j=1}^{k}Y_{ij}\right)^2-\frac{1}{mk}\left(\sum_{i=1}^{m}\sum_{j=1}^{k}Y_{ij}\right)^2$$

记

$$P=\frac{1}{mk}\left(\sum_{i=1}^{m}\sum_{j=1}^{k}Y_{ij}\right)^2$$

$$Q=\frac{1}{k}\sum_{i=1}^{m}\left(\sum_{j=1}^{k}Y_{ij}\right)^2$$

$$R=\sum_{i=1}^{m}\sum_{j=1}^{k}Y_{ij}^2$$

则

$$S_A^2=Q-P,\quad S_e^2=R-Q$$

具体计算P、Q、R这三个数时，可利用下述计算表格。

水平 数据 重复序号	1	2	…	m	
1	y_{11}	y_{21}	…	y_{m1}	
2	y_{12}	y_{22}	…	y_{m2}	
⋮	⋮	⋮		⋮	
k	y_{1k}	y_{2k}	…	y_{mk}	左边各数相加
各数相加	$\sum_{j=1}^{k}y_{1j}$	$\sum_{j=1}^{k}y_{2j}$	…	$\sum_{j=1}^{k}y_{mj}$	$\sum_{i=1}^{m}\sum_{j=1}^{k}y_{ij}$
(各数相加)^2	$\left(\sum_{j=1}^{k}y_{1j}\right)^2$	$\left(\sum_{j=1}^{k}y_{2j}\right)^2$	…	$\left(\sum_{j=1}^{k}y_{mj}\right)^2$	$\sum_{i=1}^{m}\left(\sum_{j=1}^{k}y_{ij}\right)^2$
(各数)^{2}相加	$\sum_{j=1}^{k}y_{1j}^2$	$\sum_{j=1}^{k}y_{2j}^2$	…	$\sum_{j=1}^{k}y_{mj}^2$	$\sum_{i=1}^{m}\sum_{j=1}^{k}y_{ij}^2$

注　当表6.1.1中的数据不便计算时，可对试验数据 y_{ij}，$i=$

1，2，…，m； j=1,2，…，k 作线性变换。

令 $y'_{ij} = \frac{1}{b}(y_{ij} - a) \qquad i = 1,2,\cdots,m; \quad j = 1,2,\cdots,k$

其中a,b是任意选定的两个实数。把新数据y'_{ij}填在上表中，按同样的格式计算，得出的F'值与按未作变换的数据F是相同的。因此，以后只要利用新数据进行计算即可。

事实上，由

$$Y'_{ij} = \frac{1}{b}(Y_{ij} - a)$$

有

$$\overline{Y} = a + b\overline{Y}'$$
$$\overline{Y}_{i\cdot} = a + b\overline{Y}'_{i\cdot}$$

故

$$S_e^2 = \sum_{i=1}^{m}\sum_{j=1}^{k}(Y_{ij} - \overline{Y}_{i\cdot})^2$$
$$= \sum_{i=1}^{m}\sum_{j=1}^{k}(a + bY'_{ij} - a - b\overline{Y}'_{i\cdot})^2$$
$$= b^2\sum_{i=1}^{m}\sum_{j=1}^{k}(Y'_{ij} - \overline{Y}'_{i\cdot})^2$$
$$= b^2 S'^2_e$$
$$S_A^2 = b^2 S'^2_A$$

从而

$$F = \frac{S_A^2/(m-1)}{S_e^2/(m(k-1))} = \frac{b^2S'^2_A/(m-1)}{b^2S'^2_e/m(k-1)} = F'$$

与 Y_{ij} 所取的具体值有完全一样的计算，因而对数据作线性变换，算出的F值不变。

（八）一个实例

为了熟悉上面的步骤，我们计算一个实例。

例 6.1.1　某印染厂考察染整工艺对缩水率的影响，在六种不同的染整工艺下各进行了四次试验，测得数据(%)为

试验号 \ 数据 \ 工艺	1	2	3	4	5	6
1	4.8	2.2	3.2	3.0	5.0	4.7
2	4.4	3.9	4.2	3.6	2.4	4.3
3	5.7	1.6	4.3	2.1	2.7	3.6
4	3.9	3.3	4.1	2.0	4.8	5.0

试对上述试验结果进行统计分析。

解　把上面的数据作如下的线性变换，

$$y'_{ij}=10(y_{ij}-4.0)$$

$$i=1,2,\cdots,6;\quad j=1,2,3,4.$$

将新数据 y'_{ij} 列表运算得

表

	1	2	3	4	5	6	
1	8	-18	-8	-10	10	7	
2	4	-1	2	-4	-16	3	
3	17	-24	3	-19	-13	-4	
4	-1	-7	1	-20	8	10	左边各数相加
各数相加	28	-50	-2	-53	-11	16	-72
(各数相加)^2	784	2500	4	2809	121	256	6474
(各数)^{2}相加	370	950	78	877	589	174	3038

1　显著性检验($m=6$，$k=4$)

$$P=\frac{1}{mk}\left(\sum_{i=1}^{m}\sum_{j=1}^{k}y'_{ij}\right)^2=\frac{1}{24}(-72)^2=216$$

$$Q = \frac{1}{k}\sum_{i=1}^{m}\left(\sum_{j=1}^{k}y'_{ij}\right)^2 = \frac{1}{4}6474 = 1618.5$$

$$R = \sum_{i=1}^{m}\sum_{j=1}^{k}y'^2_{ij} = 3038$$

从而 $S_A^2 = Q - P = 1618.5 - 216 = 1402.5,$

$S_e^2 = R - Q = 3038 - 1618.5 = 1419.5.$

于是统计量 F 所取的值为

$$\frac{S_A^2/(6-1)}{S_e^2/6(4-1)} = \frac{1402.5/5}{1419.5/18} = 3.56$$

给定$\alpha = 0.05$，查 F 分布表得$F_{0.05}(5,18) = 2.77$，现在$3.56 > 2.77$，所以，不同染整工艺对缩水率的影响是显著的。

2　各种工艺条件下缩水率的估计和最优工艺条件的选取

因为工艺条件的影响是显著的，所以工程平均的点估计为

$$\hat{\mu}_i = \hat{\mu} + \hat{\alpha}_i \qquad i = 1,2,\cdots,6$$

即

$$\hat{\mu}_i = \overline{y}_{i\cdot} = a + b\overline{y}'_{i\cdot}$$

由上表知

$$\overline{y}'_{1\cdot} = \frac{1}{4}\times 28 = 7,$$

$$\overline{y}'_{2\cdot} = \frac{1}{4}\times(-50) = -12.5,$$

$$\overline{y}'_{3\cdot} = \frac{1}{4}\times(-2) = -0.5,$$

$$\overline{y}'_{4\cdot} = \frac{1}{4}\times(-53) = -13.25,$$

$$\overline{y}'_{5\cdot} = \frac{1}{4}\times(-11) = -2.75,$$

$$\overline{y}'_{6\cdot} = \frac{1}{4}\times 16 = 4.$$

于是

$$\hat{\mu}_1 = 4.0 + \frac{1}{10} \times 7 = 4.7$$

$$\hat{\mu}_2 = 4.0 + \frac{1}{10} \times (-12.5) = 2.75$$

$$\hat{\mu}_3 = 4.0 + \frac{1}{10} \times (-0.5) = 3.95$$

$$\hat{\mu}_4 = 4.0 + \frac{1}{10} \times (-13.25) = 2.675$$

$$\hat{\mu}_5 = 4.0 + \frac{1}{10} \times (-2.75) = 3.725$$

$$\hat{\mu}_6 = 4.0 + \frac{1}{10} \times 4 = 4.4$$

可见，第四种工艺条件下缩水率最低，它就是最优的染整工艺。

3．各工艺条件下缩水率的区间估计

$$s_e^2 = b^2 s_e'^2 = 100^{-1} \times s_e'^2 = 14.195$$

给定$\alpha = 0.05$，查F表得$F_{0.05}(1,18) = 4.41$，算出

$$\delta = \sqrt{\frac{s_e^2 F_\alpha}{mk^2 - mk}} = \sqrt{4.41 \times \frac{14.195}{4 \times 18}} = 0.9324$$

于是，在置信度 0.95 下，各工艺条件下的缩水率的理论值应在$(\overline{y}_{i\cdot} - \delta, \overline{y}_{i\cdot} + \delta)$之内，$(i = 1, 2, \cdots, 6)$，特别地，如果采用第四种工艺，置信区间为

$$(2.675 - 0.9324,\ 2.675 + 0.9324) = (1.7426,\ 3.6074)$$

§6.2 两因子方差分析

上一节讨论了单因子试验的方差分析，这是最简单的情况。在实际生活中，更多的是两因子试验与多因子试验。这时，影响试验结果的因素往往为两个以上。我们不仅要判断各因素的影响是否显著，还要考虑因素间的交互作用的影响。这是单因子试验所没有

的．多因子的方差分析，就方法来说与两因子方差分析是一致的，区别仅在于具体的解法更加繁杂而已．因此，我们着重介绍两因子方差分析．

在两因子方差分析中，只有在每个因素的各种水平下进行了重复试验时，才能检验是否有交互作用．下面，我们给出两因子试验的数学描述．

（一）数学模型

考察两个因子 A、B 对试验结果的影响．在A的 n 个水平与 B 的m个水平的各种可能搭配下各重复地做了 l 次试验，数据如下：

表6.2.1　两因子方差分析原始数据表

数据 (A的水平 \ B的水平)	B_1	B_2	……	B_m
A_1	$Y_{111},\cdots,Y_{11l}$	$Y_{121},\cdots,Y_{12l}$	……	$Y_{1m1},\cdots,Y_{1ml}$
A_2	$Y_{211},\cdots,Y_{21l}$	$Y_{221},\cdots,Y_{22l}$	……	$Y_{2m1},\cdots,Y_{2ml}$
⋮	⋮	⋮		⋮
A_n	$Y_{n11},\cdots,Y_{n1l}$	$Y_{n21},\cdots,Y_{n2l}$	……	$Y_{nm1},\cdots,Y_{nml}$

对上表的数据进行统计分析，主要应回答下面一些问题：

1　因子A、B中，哪个对试验结果的影响是显著的；

2　决定各因子水平的最佳搭配，即选择“最优”生产条件；

3　对一般平均及 A、B 各水平的效应作估计，并由显著性检验的结果对各因子水平搭配下的工程平均作出点估计和区间估计．

由于在各次试验时均有随机误差的干扰，因此，表 6.2.1 的数据

$$Y_{ijk}=\mu_{ij}+\varepsilon_{ijk},\quad i=1,2,\cdots,n,\quad j=1,2,\cdots,m,\quad k=1,2,\cdots,l \tag{6.2.1}$$

其中μ_{ij}表示在水平组合A_i，B_j下的工程平均，ε_{ijk}表示在水平组

合 A_i, B_j 下，第 k 次重复试验时的试验误差（即随机误差）。基于和讨论单因子试验时同样的考虑，我们可以假定 ε_{ijk} 是 $m \cdot n \cdot l$ 个相互独立，同正态 $N(0, \sigma^2)$ 的随机变量，$i=1,2,\cdots,n, j=1,2,\cdots,m, k=1,2,\cdots,l$。

令

$$\mu = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\mu_{ij} \tag{6.2.2}$$

$$\mu_{i\cdot} = \frac{1}{m}\sum_{j=1}^{m}\mu_{ij} \tag{6.2.3}$$

$$\mu_{\cdot j} = \frac{1}{n}\sum_{i=1}^{n}\mu_{ij} \tag{6.2.4}$$

称 μ 为一般平均。

以后，我们把分别反映 A_i 和 B_j 对试验结果作用大小的量

$$\alpha_i = \mu_{i\cdot} - \mu, \quad i=1,2,\cdots,n$$
$$\beta_j = \mu_{\cdot j} - \mu, \quad j=1,2,\cdots,m$$

分别称为 A_i 与 B_j 的水平效应，对 α_i，β_j 显然有：

$$\sum_{i=1}^{n}\alpha_i = 0 \tag{6.2.5}$$

$$\sum_{j=1}^{m}\beta_j = 0 \tag{6.2.6}$$

为了分析因子各水平的作用，把水平组合 A_i, B_j 下的工程平均 μ_{ij} 表成

$$\begin{aligned}\mu_{ij} &= \mu + (\mu_{i\cdot} - \mu) + (\mu_{\cdot j} - \mu) + (\mu_{ij} - \mu_{i\cdot} - \mu_{\cdot j} + \mu) \\ &= \mu + \alpha_i + \beta_j + (\mu_{ij} - \mu_{i\cdot} - \mu_{\cdot j} + \mu) \qquad (6.2.7) \\ & \qquad i=1,2,\cdots,n, \quad j=1,2,\cdots,m\end{aligned}$$

记 $(\alpha\beta)_{ij} = \mu_{ij} - \mu_{i\cdot} - \mu_{\cdot j} + \mu$，$i=1,2,\cdots,n, j=1,2,\cdots,m$。把 $(\alpha\beta)_{ij}$ 称为 A 的第 i 个水平 A_i 与 B 的第 j 个水平 B_j 的交互作用效应，则

$$\sum_{i=1}^{m}(\alpha\beta)_{ij} = \sum_{j=1}^{m}(\mu_{ij} - \mu_{i\cdot} - \mu_{\cdot j} + \mu) = 0 \qquad (6.2.8)$$

$$i = 1,2,\cdots,n$$

$$\sum_{i=1}^{n}(\alpha\beta)_{ij} = \sum_{i=1}^{n}(\mu_{ij} - \mu_{i\cdot} - \mu_{\cdot j} + \mu) = 0 \qquad (6.2.9)$$

$$j = 1,2,\cdots,m$$

在处理两因子试验问题时，往往把A与B各水平搭配下的交互作用效应设想为某一个因子的作用，这个因子记为$A\times B$，称为A与B的交互作用。

总之，

$$Y_{ijk} = \mu_{ij} + \varepsilon_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \qquad (6.2.10)$$

$$i = 1,2,\cdots,n, \quad j = 1,2,\cdots,m, \quad k = 1,2,\cdots,l$$

称为两因子试验的数学模型，式中诸α_i，诸β_j，诸$(\alpha\beta)_{ij}$分别满足(6.2.5),(6.2.6),(6.2.8)，(6.2.9) 式。

（二）参数估计

由于ε_{ijk}是$n\cdot m\cdot l$个相互独立且具有相同分布$N(0, \sigma^2)$的随机变量，由 (6.2.10) 可知，对固定的i和j．

$$Y_{ijk} \sim N(\mu_{ij}, \sigma^2)$$

$$i = 1,2,\cdots,n, \quad j = 1,2,\cdots,m, \quad k = 1,2,\cdots,l$$

因此，我们用子样均值

$$\overline{Y}_{ij\cdot} = \frac{1}{l}\sum_{k=1}^{l} Y_{ijk} \qquad (6.2.11)$$

去估计总体均值μ_{ij}，即$\hat{\mu}_{ij} = \overline{Y}_{ij\cdot}$，$i = 1,2,\cdots,n, j = 1,2,\cdots,m$。

由

$$E(\hat{\mu}_{ij}) = E\left(\frac{1}{l}\sum_{k=1}^{l} Y_{ijk}\right) = \frac{1}{l}\sum_{k=1}^{l} E(Y_{ijk})$$

$$= \frac{1}{l}\sum_{k=1}^{l}\mu_{ij} = \mu_{ij}$$

知，$\hat{\mu}_{ij}$是μ_{ij}的无偏估计。

因为 $\mu = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\mu_{ij}$，因而可令

$$\hat{\mu} = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\hat{\mu}_{ij} = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\overline{Y}_{ij\cdot}$$

$$= \frac{1}{nml}\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{l}Y_{ijk} = \overline{Y}$$

为μ的估计量，且有$E(\hat{\mu}) = \mu$。

令

$$\hat{\mu}_{i\cdot} = \frac{1}{m}\sum_{j=1}^{m}\hat{\mu}_{ij} = \frac{1}{m}\sum_{j=1}^{m}\overline{Y}_{ij\cdot} = \frac{1}{ml}\sum_{j=1}^{m}\sum_{k=1}^{l}Y_{ijk} = \overline{Y}_{i\cdot\cdot}$$

$$i = 1, 2, \cdots, n \tag{6.2.12}$$

$$\hat{\mu}_{\cdot j} = \frac{1}{n}\sum_{i=1}^{n}\hat{\mu}_{ij} = \frac{1}{n}\sum_{i=1}^{n}\overline{Y}_{ij\cdot} = \frac{1}{nl}\sum_{i=1}^{n}\sum_{k=1}^{l}Y_{ijk}$$

$$= \overline{Y}_{\cdot j\cdot} \qquad j = 1, 2, \cdots, m \tag{6.2.13}$$

$$\hat{\alpha}_i = \hat{\mu}_{i\cdot} - \hat{\mu} = \overline{Y}_{i\cdot\cdot} - \overline{Y}, \tag{6.2.14}$$

$$\hat{\beta}_j = \hat{\mu}_{\cdot j} - \hat{\mu} = \overline{Y}_{\cdot j\cdot} - \overline{Y} \tag{6.2.15}$$

$$\widehat{(\alpha\beta)}_{ij} = \hat{\mu}_{ij} - \hat{\mu}_{i\cdot} - \hat{\mu}_{\cdot j} + \hat{\mu}$$

$$= \overline{Y}_{ij\cdot} - \overline{Y}_{i\cdot\cdot} - \overline{Y}_{\cdot j\cdot} + \overline{Y} \tag{6.2.16}$$

易知，它们分别是$\mu_{i\cdot}$，$\mu_{\cdot j}$，α_i，β_j，$(\alpha\beta)_{ij}$的无偏估计，且满足

$$\sum_{i=1}^{n}\hat{\alpha}_i = 0$$

$$\sum_{j=1}^{m}\hat{\beta}_j = 0$$

$$\sum_{j=1}^{m}(\widehat{\alpha\beta})_{ij}=0 \qquad i=1,2,\cdots,n$$

$$\sum_{i=1}^{n}(\widehat{\alpha\beta})_{ij}=0 \qquad j=1,2,\cdots,m$$

由此，

$$\begin{aligned} Y_{ijk} &= \overline{Y}+(\overline{Y}_{i\cdot\cdot}-\overline{Y})+(\overline{Y}_{\cdot j\cdot}-\overline{Y})+ \\ &\quad (\overline{Y}_{ij\cdot}-\overline{Y}_{i\cdot\cdot}-\overline{Y}_{\cdot j\cdot}+\overline{Y})+Y_{ijk}-\overline{Y}_{ij\cdot} \\ &= \hat{\mu}+\hat{\alpha}_i+\hat{\beta}_j+(\widehat{\alpha\beta})_{ij}+e_{ijk} \end{aligned} \tag{6.2.17}$$

$$i=1,2,\cdots,n; \quad j=1,2,\cdots,m; \quad k=1,2,\cdots,l$$

式中 $e_{ijk}=Y_{ijk}-\overline{Y}_{ij\cdot}$ 是数据中反映误差的部分。

（三）显著性检验

与单因子试验时的作法一样，我们先求总的偏差平方和。由(6.2.17)式得

$$Y_{ijk}-\hat{\mu}=\hat{\alpha}_i+\hat{\beta}_j+(\widehat{\alpha\beta})_{ij}+e_{ijk}$$

上式两端平方得：

$$\begin{aligned} (Y_{ijk}-\hat{\mu})^2 &= \hat{\alpha}_i^2+\hat{\beta}_j^2+(\widehat{\alpha\beta})_{ij}^2+e_{ijk}^2+2\hat{\alpha}_i\hat{\beta}_j+ \\ &\quad +2\hat{\alpha}_i(\widehat{\alpha\beta})_{ij}+2\hat{\alpha}_i e_{ijk}+2\hat{\beta}_j(\widehat{\alpha\beta})_{ij}+2\hat{\beta}_j e_{ijk} \\ &\quad +2(\widehat{\alpha\beta})_{ij}e_{ijk} \end{aligned}$$

对一切i,j,k求和，并注意到所有交叉项求和均为0，例如，

$$\begin{aligned} \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{l}\hat{\alpha}_i\hat{\beta}_j &= l\sum_{i=1}^{n}\sum_{j=1}^{m}(\overline{Y}_{i\cdot\cdot}-\overline{Y})(\overline{Y}_{\cdot j\cdot}-\overline{Y}) \\ &= l\sum_{i=1}^{n}(\overline{Y}_{i\cdot\cdot}-\overline{Y})\sum_{j=1}^{m}(\overline{Y}_{\cdot j\cdot}-\overline{Y})=0 \end{aligned}$$

便得总平方和的分解式，

$$S_T^2=S_A^2+S_B^2+S_{A\times B}^2+S_e^2 \tag{6.2.18}$$

其中，

$$S_T^2=\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{l}(Y_{ijk}-\overline{Y})^2$$

$$S_A^2 = ml\sum_{i=1}^{n}\hat{\alpha}_i^2 = ml\sum_{i=1}^{n}(\overline{Y}_{i..} - \overline{Y})^2,$$

$$S_B^2 = nl\sum_{j=1}^{m}\hat{\beta}_j^2 = nl\sum_{j=1}^{m}(\overline{Y}_{.j.} - \overline{Y})^2,$$

$$S_{A\times B}^2 = l\sum_{i=1}^{n}\sum_{j=1}^{m}(\widehat{\alpha\beta})_{ij}^2$$

$$= l\sum_{i=1}^{n}\sum_{j=1}^{m}(\overline{Y}_{ij.} - \overline{Y}_{i..} - \overline{Y}_{.j.} + \overline{Y})^2$$

$$S_e^2 = \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{l}(Y_{ijk} - \overline{Y}_{ij.})^2$$

它们的意义是，数据的总波动S_T^2可分解为四部分，第一部分S_A^2用来反映因子A引起的波动，第二部分S_B^2用来反映因子B引起的波动，第三部分$S_{A\times B}^2$用来反映交互作用$A\times B$引起的波动，第四部分S_e^2则完全是由于随机误差所引起的，分别称它们为A的偏差平方和，B的偏差平方和，$A\times B$的偏差平方和与误差的偏差平方和。

为了进行显著性检验的需要，计算

$$ES_e^2 = \sum_{i=1}^{n}\sum_{j=1}^{m}E\left(\sum_{k=1}^{l}(Y_{ijk} - \overline{Y}_{ij.})^2\right) = nm(l-1)\sigma^2$$

即

$$E\left(\frac{S_e^2}{nm(l-1)}\right) = \sigma^2 \tag{6.2.19}$$

上式表明，

$$\hat{\sigma}^2 = \frac{S_e^2}{nm(l-1)}$$

是σ^2的无偏估计。由于σ^2完全刻划了误差ε_{ijk}的性质，所以$\hat{\sigma}^2$实质上就是对误差ε_{ijk}的一种估计。类似于（6.1.8）的计算，有

$$ES_A^2 = (n-1)\sigma^2 + ml\sum_{i=1}^{n}\alpha_i^2 \tag{6.2.20}$$

$$ES_B^2=(m-1)\sigma^2+nl\sum_{j=1}^{m}\beta_j^2 \tag{6.2.21}$$

$$ES_{A\times B}^2=(n-1)(m-1)\sigma^2+\sum_{i=1}^{n}\sum_{j=1}^{m}(\alpha\beta)_{ij}^2 \tag{6.2.22}$$

由这些式子可知，检验A的作用，B的作用及A,B的交互作用$A\times B$实际上是分别检验假设

H_{01}：$\alpha_1=\alpha_2=\cdots=\alpha_n=0$，

H_{02}：$\beta_1=\beta_2=\cdots=\beta_m=0$，

H_{03}：$(\alpha\beta)_{ij}=0$，$i=1,2,\cdots,n$，$j=1,2,\cdots,m$

是否成立。

利用柯赫伦定理可以证明（证明过程略）如下一些结果：

1　若H_{01}，H_{02}，H_{03}同时成立（注意，这时$Y_{ijk}\sim N(\mu,\sigma^2)$）则

$$\frac{S_A^2}{\sigma^2},\quad \frac{S_B^2}{\sigma^2},\quad \frac{S_{A\times B}^2}{\sigma^2},\quad \frac{S_e^2}{\sigma^2}$$

相互独立，且分别为$\chi^2(n-1)$，$\chi^2(m-1)$，$\chi^2((n-1)(m-1))$和$\chi^2(nm(l-1))$变量；

2　若H_{01}成立，则$\frac{S_A^2}{\sigma^2}$是$\chi^2(n-1)$变量且与S_e^2相互独立；

3　若H_{02}成立，则$\frac{S_B^2}{\sigma^2}$是$\chi^2(m-1)$变量且与S_e^2相互独立；

4　若H_{03}成立，则$\frac{S_{A\times B}^2}{\sigma^2}$是$\chi^2[(m-1)(n-1)]$变量且与$S_e^2$相互独立；

5　不论H_{01}，H_{02}，H_{03}成立与否，$\frac{1}{\sigma^2}S_e^2$是$\chi^2[nm(l-1)]$变量。

仿照单因子方差分析的作法，可选用F变量来作为检验假设H_{01}，H_{02}，H_{03}的统计量

当H_{01}成立时，

$$F_A=\frac{S_A^2/(n-1)}{S_e^2/nm(l-1)}\sim F(n-1,\ nm(l-1)) \qquad (6.2.23)$$

当 H_{02} 成立时，

$$F_B=\frac{\dfrac{S_B^2}{m-1}}{\dfrac{S_e^2}{nm(l-1)}}\sim F(m-1,\ nm(l-1)) \qquad (6.2.24)$$

当 H_{03} 成立时

$$F_{A\times B}=\frac{\dfrac{S_{A\times B}^2}{(n-1)(m-1)}}{\dfrac{S_e^2}{nm(l-1)}}\sim F[(n-1)(m-1),nm(l-1)] \qquad (6.2.25)$$

由此，当给定显著水平 α 以后，便可以通过计算 F_A，F_B，$F_{A\times B}$ 的数值，与 F 的临界值作比较来判断因子A、B及交互作用 $A\times B$的显著性。

（四）工程平均的估计

在显著性检验之后，使我们有可能根据检验的结果对各种试验条件下的工程平均进行新的估计。本段的目的是给出各种条件下工程平均的点估计与区间估计，即对试验结果进行“予测”。

如果显著性检验的结果是因子A不显著（即 H_{01} 成立），而因子B和交互作用$A\times B$是显著的。则（6.2.7）变成

$$\mu_{ij}=\mu+\beta_j+(\alpha\beta)_{ij} \qquad i=1,2,\cdots,n,\ j=1,2,\cdots,m$$

这时，工程平均的点估计应是

$$\hat{\mu}_{ij}=\hat{\mu}+\hat{\beta}_j+\widehat{(\alpha\beta)}_{ij}, \qquad i=1,2,\cdots,n,j=1,2,\cdots,m \qquad (6.2.26)$$

一般说来，工程平均的点估计，可归纳成为下述用文字叙述的公式：

工程平均的估计值＝一般平均的估计值＋显著因子相应水平效应的估计值　　(6.2.27)

下面再求工程平均的区间估计，如果因子A是不显著的，则由

(6.2.20) 式有

$$ES_A^2=(n-1)\sigma^2$$

这表明，在因子A的偏差平方和S_A^2中，因子A的水平不同所引起的波动为0，S_A^2反映的全部是误差所引起的波动。因此，应该把S_A^2也并入误差的偏差平方和，且以

$$\widetilde{S}_e^2=S_A^2+S_e^2$$

来代替原来的S_e^2作为误差的偏差平方和。因为S_A^2与S_e^2相互独立，由χ^2分布的可加性知

$$\frac{\widetilde{S}_e^2}{\sigma^2}\sim\chi^2[n-1+nm(l-1)]$$

为方便起见，用f_A记S_A^2的秩（也称A的自由度），用f_e记S_e^2的秩，其它偏差平方和也采用类似的记号。

数学上可以证明，当假设H_{01}成立时，$\hat{\mu}_{ij}=\hat{\mu}+\hat{\beta}_j+(\widehat{\alpha\beta})_{ij}$与$\widetilde{S}_e^2$相互独立（对一切$i$，$j$），且

$$E(\hat{\mu}_{ij}-\mu_{ij})^2=D\hat{\mu}_{ij}=\frac{1+f_B+f_{A\times B}}{nml}\sigma^2$$

记

$$n_e=\frac{nml}{1+f_B+f_{A\times B}}$$

则有

$$\frac{n_e(\hat{\mu}_{ij}-\mu_{ij})^2}{\dfrac{\widetilde{S}_e^2}{\widetilde{f}_e}}\sim F(1,\widetilde{f}_e) \qquad (6.2.28)$$

对于显著性检验时所出现的另外一些情况(这时，不外乎是有些因子显著，有些不显著），有类似于上式的结果。这时，只要分别把$\widetilde{S}_e^2$，$\widetilde{f}_e^2$和n_e的表达式用文字表成*

---

* 前面把$A\times B$也看作一个因子，不过交互作用的自由度与各个因子的自由度的计算方法不同，因子自由度等于其水平数减1，两因子的交互作用的自由度等于两个因子的自由度的乘积。

$$
\begin{cases}
\widetilde{S}_e^2 = S_e^2 + \text{不显著因子的偏差平方和}, \\
\widetilde{f}_e = f_e + \text{不显著因子的自由度之和}, \\
n_e = \dfrac{\text{试验总次数}}{1+\text{显著因子的自由度之和}}
\end{cases}
\quad (6.2.29)
$$

即可。当然（6.2.26）中的$\hat{\mu}_{ij}$也要用以文字表达的公式(6.2.27)来计算。由（6.2.28）知，给定置信度$1-\alpha$，就可由F分布表查出临界值 $F_\alpha(1, \widetilde{f}_e)$，使得

$$
P\left(\frac{n_e(\hat{\mu}_{ij}-\mu_{ij})^2}{\widetilde{S}_e^2/\widetilde{f}_e} < F_\alpha\right) = 1-\alpha
$$

即

$$
P\left(\hat{\mu}_{ij} - \sqrt{\frac{F_\alpha \cdot \widetilde{S}_e^2}{n_e \widetilde{f}_e}} < \mu_{ij} < \hat{\mu}_{ij} + \sqrt{\frac{F_\alpha \cdot \widetilde{S}_e^2}{n_e \widetilde{f}_e}}\right)
$$
$$
= 1-\alpha \quad (6.2.30)
$$

上式表明，对给定的α，通过试验数据算出$\hat{\mu}_{ij}$和$\widetilde{S}_e^2$后，就有$1-\alpha$的把握“预测”工程平均落在置信区间（$\hat{\mu}_{ij}-\delta, \hat{\mu}_{ij}+\delta$）中，这里

$$
\delta = \sqrt{\frac{F_\alpha \cdot \widetilde{S}_e^2}{n_e \widetilde{f}_e}} \quad (6.2.31)
$$

（五）计算方法和例

与§6.1的计算方法一样，统计分析的计算步骤如下：

1　适当选择常数a，$b \neq 0$，对原试验数据作变换

$$
y'_{ijk} = \frac{1}{b}(y_{ijk} - a)
$$

$$
i = 1, 2, \cdots, n, \quad j = 1, 2, \cdots, m, \quad k = 1, 2, \cdots, l.
$$

使新数据便于计算。（如果原数据算起来很简便，这一步就没有必要了。）

2 列表计算

因子 水平 试验号	A	B	数据 y'_{ijk}	合计
1	1	1	$y'_{111}\cdots y'_{11l}$	$\sum_{k=1}^{l} y'_{11k}$
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
m	1	m	$y'_{1m1}\cdots y'_{1ml}$	$\sum_{k=1}^{l} y'_{1mk}$
$m+1$	2	1	$y'_{211}\cdots y'_{21l}$	$\sum_{k=1}^{l} y'_{21k}$
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
$2m$	2	m	$y'_{2m1}\cdots y'_{2ml}$	$\sum_{k=1}^{l} y'_{2mk}$
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
$(n-1)m+1$	n	1	$y'_{n11}\cdots y'_{n1l}$	$\sum_{k=1}^{l} y'_{n1k}$
⋮	⋮	⋮	⋮	⋮
nm	n	m	$y'_{nm1}\cdots y'_{nml}$	$\sum_{k=1}^{l} y'_{nmk}$
各列相应于"1"的"合计"之和	$\sum_{j=1}^{m}\sum_{k=1}^{l} y'_{1jk}$	$\sum_{i=1}^{n}\sum_{k=1}^{l} y'_{i1k}$		
⋮	⋮	⋮		
各列相应于"m"的"合计"之和	⋮	$\sum_{i=1}^{n}\sum_{k=1}^{l} y'_{imk}$		
⋮				
各列相应于"n"的"合计"之和	$\sum_{j=1}^{m}\sum_{k=1}^{l} y'_{njk}$			

由上表算出的结果有

$$\sum_{k=1}^{l} y'_{ijk} \quad i=1,2,\cdots,n, \quad j=1,2,\cdots,m$$

$$\sum_{j=1}^{m}\sum_{k=1}^{l} y'_{ijk} \quad i=1,2,\cdots,n$$

$$\sum_{i=1}^{n}\sum_{k=1}^{l} y'_{ijk} \quad i=1,2,\cdots,m$$

3 把上述计算结果代入公式

$$\hat{\mu}'_{ij}=\frac{1}{l}\sum_{k=1}^{l} y'_{ijk} \quad i=1,2,\cdots,n, \; j=1,2,\cdots,m$$

$$\hat{\mu}'_{i\cdot}=\frac{1}{ml}\sum_{j=1}^{m}\sum_{k=1}^{l} y'_{ijk} \quad i=1,2,\cdots,n$$

$$\hat{\mu}'_{\cdot j}=\frac{1}{nl}\sum_{i=1}^{n}\sum_{k=1}^{l} y'_{ijk} \quad j=1,2,\cdots,m$$

求出$\hat{\mu}'_{ij}$，$\hat{\mu}'_{i\cdot}$，$\hat{\mu}'_{\cdot j}$等，并由它们进一步求出一般平均与各因子效应的估计值$\hat{\mu}'$，$\hat{\alpha}'_i$，$\hat{\beta}'_j$，$(\widehat{\alpha\beta})'_{ij}$，$i=1,2,\cdots,n$，$j=1,2,\cdots,m$。通过这些数，再由变换公式就可算出$\hat{\mu}$，$\hat{\alpha}_i$，$\hat{\beta}_j$，$(\widehat{\alpha\beta})_{ij}$，$i=1,2,\cdots,n$，$j=1,2,\cdots,m$。

由（6.1.14）知

$$S'^2_T=\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{l}(y'_{ijk}-\overline{y}')^2$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{l} y'^2_{ijk}-CT'$$

其中

$$CT'=\frac{1}{mnl}\left(\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{l} y'_{ijk}\right)^2$$

类似地有

$$S_A'^2 = \frac{1}{ml}\sum_{i=1}^{n}\left(\sum_{j=1}^{m}\sum_{k=1}^{l}y'_{ijk}\right)^2 - CT'$$

$$S_B'^2 = \frac{1}{nl}\sum_{j=1}^{m}\left(\sum_{i=1}^{n}\sum_{k=1}^{l}y'_{ijk}\right)^2 - CT'$$

$$S_{A\times B}'^2 = \frac{1}{l}\sum_{i=1}^{n}\sum_{j=1}^{m}\left(\sum_{k=1}^{l}y'_{ijk}\right)^2 - \frac{1}{ml}\sum_{i=1}^{n}\left(\sum_{j=1}^{m}\sum_{k=1}^{l}y'_{ijk}\right)^2$$

$$- \frac{1}{nl}\sum_{j=1}^{m}\left(\sum_{i=1}^{n}\sum_{k=1}^{l}y'_{ijk}\right)^2 + CT'$$

$$S_e'^2 = \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{l}y'^2_{ijk} - \frac{1}{l}\sum_{i=1}^{n}\sum_{j=1}^{m}\left(\sum_{k=1}^{l}y'_{ijk}\right)^2$$

把上表算出的结果代入上述各式，即可算出$S_T'^2$，$S_A'^2$，$S_B'^2$，$S_{A\times B}'^2$，$S_e'^2$，将这些数值再乘以b^2就得出相应与原始数据的S_T^2，S_A^2，S_B^2，$S_{A\times B}^2$，S_e^2。由于变换前后算出的F_A，F_B，$F_{A\times B}$的数值不变，因此，进行显著性检验时，直接对变换后的数据运算就可以了。

例6.2.1　硬橡胶带抗牵拉强度试验

抗牵拉强度是硬橡胶带的一项重要的性能指标。试验考察两个因子对抗牵拉强度的影响：

A（硫化时间）A_1　40″，A_2　60″

B（促进剂种类）B_1　甲，B_2　乙，B_3　丙

在A与B六种水平组合下，各重复作了两次试验，测得数据（单位：kg/cm²）为

数据 B \ A	B_1	B_2	B_3
A_1	390 380	440 420	370 350
A_2	390 410	450 430	370 380

对上述试验数据进行统计分析。

解　分下列几步进行

1　对原始数据作变换，取$a=400$，$b=10$，令

$$y'_{ijk}=\frac{1}{10}(y_{ijk}-400)$$

$$i=1,2,\quad j=1,2,3,\quad k=1,2$$

简化后的数据为

A \ B（数据）	B_1	B_2	B_3
A_1	-1 -2	4 2	-3 -5
A_2	-1 1	5 3	-3 -2

2　将简化后的数据列表进行计算

	A	B	数据 y'_{ijk}		合计
1	1	1	-1	-2	-3
2	1	2	4	2	6
3	1	3	-3	-5	-8
4	2	1	-1	1	0
5	2	2	5	3	8
6	2	3	-3	-2	-5
各列相应于“1”的“合计”之和	-5	-3			
各列相应于“2”的“合计”之和	3	14			
各列相应于“3”的“合计”之和		-13			

将这个表的“合计”一栏相加，得数据的总和为-2。

3　统计分析

（1）显著性检验。将表中计算出的结果代入CT'，S'^2_T，S'^2_A，

$S_B'^2$，$S_{A\times B}'^2$，$S_e'^2$中得

$$CT' = \frac{1}{12}\times(-2)^2 = \frac{1}{3}$$

$$S_T'^2 = (-1)^2+(-2)^2+4^2+2^2+(-3)^2+(-5)^2+(-1)^2 + 1^2+5^2+3^2+(-3)^2+(-2)^2 - \frac{1}{3} = \frac{323}{3}$$

$$S_A'^2 = \frac{1}{3\times 2}[(-5)^2+3^2] - \frac{1}{3} = \frac{16}{3}$$

$$S_B'^2 = \frac{1}{2\times 2}[(-3)^2+14^2+(-13)^2] - \frac{1}{3} = \frac{559}{6}$$

$$S_{A\times B}'^2 = \frac{1}{2}[(-3)^2+6^2+(-8)^2+0^2+8^2+(-5)^2] - \frac{16}{3} - \frac{559}{6} - \frac{1}{3} = \frac{1}{6}$$

$$S_e'^2 = 9$$

于是得

$$F_A = \frac{\frac{16}{3}/1}{9/2\times 3\times 1} \doteq 3.56$$

$$F_B = \frac{\frac{559}{6}/2}{9/6} \doteq 31.06$$

$$F_{A\times B} = \frac{\frac{1}{6}/2}{9/6} \doteq 0.06$$

查 F 表可知 $F_{0.05}(1,6)=5.99$，$F_{0.01}(2,6)=10.92$、这样，就知道因子 A 和交互作用 $A\times B$ 都是不显著的，而因子 B 则高度显著。

（2）最优生产条件的选取．由于因子 B 高度显著，故必须注意选取 B 的水平使抗牵拉强度最大．为此，计算 B 的三个效应：

$$\hat{\beta}_1 = 10\times\left(-\frac{3}{4}+\frac{2}{12}\right) = -5.8$$

$$\hat{\beta}_2=10\times\left(\frac{14}{4}+\frac{2}{12}\right)=36.7$$

$$\hat{\beta}_3=10\times\left(-\frac{13}{4}+\frac{2}{12}\right)=-30.8$$

由此看出，以$\hat{\beta}_2$为最大. 这说明乙种促进剂对提高抗牵拉强度效果最好. 又A是不显著因子，它的水平可以任取. 从工艺上考虑，硫化时间越短，生产效率越高. 所以取A的一水平为宜. 这样选出的最优生产条件是A_1B_2，即反应时间40″，用乙种促进剂.

（3）工程平均的估计. 因为一般平均为

$$\hat{\mu}=a+b\hat{\mu}'=400+10\times\left(-\frac{2}{12}\right)=398.3$$

由（6.2.26）式，可求出各条件下工程平均的估计值

$$\hat{\mu}_{11}=\hat{\mu}_{21}=\hat{\mu}+\hat{\beta}_1=398.3-5.8=392.5$$
$$\hat{\mu}_{12}=\hat{\mu}_{22}=\hat{\mu}+\hat{\beta}_2=398.3+36.7=435$$
$$\hat{\mu}_{13}=\hat{\mu}_{23}=\hat{\mu}+\hat{\beta}_3=398.3-30.8=367.5$$

在求区间估计时，应把不显著因子的变动平方和并入误差，得

$$\tilde{S}_e^2=S_e^2+S_A^2+S_{A\times B}^2=\left(9+\frac{16}{3}+\frac{1}{6}\right)\times100=1450$$

$$\tilde{f}_e=f_e+f_A+f_{A\times B}=6+1+2=9$$

$$n_e=\frac{12}{1+2}=4$$

$$F_{0.05}(1,9)=5.12$$

由此算出

$$\delta=\sqrt{\frac{5.12\times1450}{9\times4}}=14.36$$

因此，在各试验条件下工程平均的置信区间为

A_1B_1和A_2B_1：（392.5－14.36，392.5＋14.36）

A_1B_2和A_2B_2：（435－14.36，435＋14.36）

A_1B_3和A_2B_3：（367.5－14.36，367.5＋14.36）

特别地，如果采用所选出的最优生产条件A_1B_2进行生产，可以

预测抗牵拉强度的理论值将在

$$435-14.36=420.64(\text{kg/cm}^2)$$

与

$$435+14.36=449.36(\text{kg/cm}^2)$$

之间。

§6.3 一元线性回归

我们从最简单的情况，即一个变量与一个变量之间的相关关系开始。

（一）最小二乘法

先看一个例子。

例 6.3.1 在考察硝酸钠 $NaNO_3$ 的可溶性的程度时，对一列不同的温度，观察它在100ml的水中溶解的 $NaNO_3$ 的重量，得观察结果如下：

温度(℃) x_i	0	4	10	15	21	29	36	51	68
$NaNO_3$ 重(g) y_i	66.7	71.0	76.3	80.6	85.7	92.9	99.4	113.6	125.1

依照表中的数据作图，则可看出，硝酸钠的溶重 y_i 随温度 x_i 的升高而增大，大体上呈现直线关系。

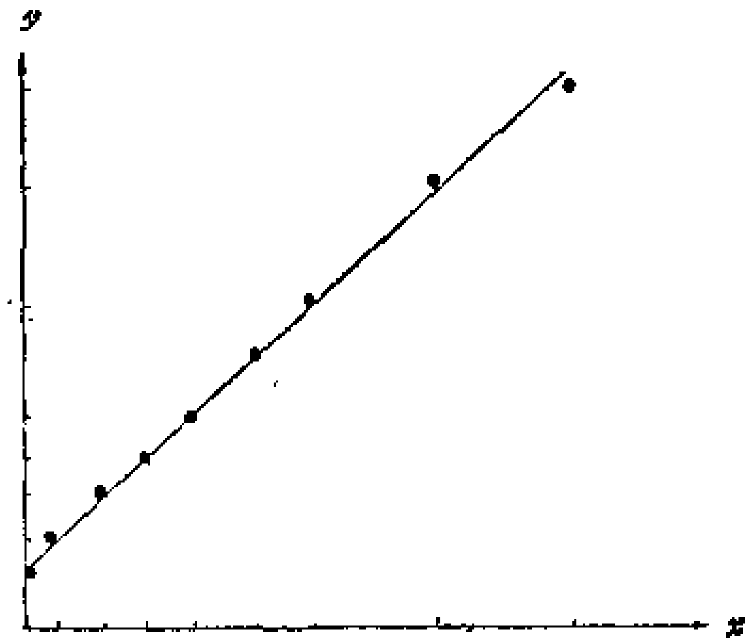
显然，我们还必须求出溶重与温度之间的定量关系式。

由图6.3.1，我们看到定量公式的形式是完全确定的，即

$$\hat{y}=\alpha+\beta x, \tag{6.3.1}$$

这里，$\hat{y}$ 是当温度取值 x 时，硝酸钠溶重的真值，它区别于带有随机误差的实际值 y。

由此，只需要确定公式（6.3.1）中的系数 α 和 β 就可以了。这里 β 通常称为回归系数。关系式 $\hat{y}=\alpha+\beta x$ 叫做回归方程。它的几何图象是平面上过（0,α）点，斜率为 β 的一条直线，称为回归



图（6.3.1）

直线．

直观上，要在平面上画出这样一条表示y和x之间关系的直线是不困难的．我们只要使该直线总的看来最“接近”这些点子就可以了．这种方法虽然简便易行，但十分粗糙，画出的直线（从而得到它的直线方程）因人而异，各不相同．在考虑多个变量之间的相关关系时，这种直观方法就行不通了．虽然有这些缺点，但它的基本思想却是可取的，即提出了一个“总的看来最接近”的原则．由此，就导出了最小二乘法．

给定平面上的n个点(x_1, y_1)，(x_2, y_2)，…，(x_n, y_n)，对于平面上的任意一条直线l

$$y = a + bx,$$

我们用数量

$$|y_i - (a + bx_i)|^2$$

来刻划点(x_i, y_i)到直线l的远近程度。当点(x_i, y_i)在直线l上$(y_i = a + bx_i)$时，有$|y_i - (a + bx_i)|^2 = 0$最小；当点(x_i, y_i)远离直线l时，$|y_i - (a + bx_i)|^2$就远远大于0．因而

$$\sum_{i=1}^{n} |y_i - (a + bx_i)|^2$$

就定量地描述了直线 l 与这 n 个点的总的接近的程度。值得注意的是，在这个量中，x_i，y_i是定数，a，b是变数，即它是a，b的二元函数

$$Q(a,b) \triangleq \sum_{i=1}^{n} [y_i - (a+bx_i)]^2 \tag{6.3.2}$$

找一条直线最接近这n个点，就是求使得$Q(a, b)$达到极小的a，b值。于是，我们的问题可提为：

要求两个数$\hat{\alpha}$，$\hat{\beta}$，使二元函数$Q(a, b)$，在$a=\hat{\alpha}$，$b=\hat{\beta}$时达到极小值。

由于$Q(a, b)$是n个平方数的和，所以"使$Q(a,b)$达到极小"的原则称为平方和（二乘方和）最小的原则，习惯上简称为最小二乘原则。依据最小二乘原则去求$\hat{\alpha}$，$\hat{\beta}$的方法叫最小二乘法。

由于$Q(a, b)$是a，b的非负的二次函数，求Q的最小解是有可能的。所要求的$\hat{\alpha}$，$\hat{\beta}$应满足联立方程

$$\begin{cases} \dfrac{\partial Q(a, b)}{\partial a} = 0 \\ \dfrac{\partial Q(a, b)}{\partial b} = 0 \end{cases}$$

由于

$$\frac{\partial Q(a, b)}{\partial a} = -2\sum_{i=1}^{n}[y_i - (a+bx_i)]$$

$$\frac{\partial Q(a, b)}{\partial b} = -2\sum_{i=1}^{n}[y_i - (a+bx_i)]x_i$$

故要求解的线性方程组为

$$\sum_{i=1}^{n}[y_i - (a+bx_i)] = 0 \tag{6.3.3}$$

$$\sum_{i=1}^{n}[y_i - (a+bx_i)]x_i = 0 \tag{6.3.4}$$

由（6.3.3）得

$$a = \frac{1}{n}\sum_{i=1}^{n} y_i - b\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) = \overline{y} - b\overline{x}$$

用（6.3.4）式减去$\overline{x}$倍（6.3.3）式得

$$\sum_{i=1}^{n}[y_i - (a + bx_i)](x_i - \overline{x}) = 0$$

把$a = \overline{y} - b\overline{x}$代入上式得

$$\sum_{i=1}^{n}[(y_i - \overline{y}) - b(x_i - \overline{x})](x_i - \overline{x}) = 0$$

$$b = \frac{\sum_{i=1}^{n}(y_i - \overline{y})(x_i - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

记

$$\hat{\alpha} = \overline{y} - \hat{\beta}\,\overline{x} \tag{6.3.5}$$

$$\hat{\beta} = \frac{\sum_{i=1}^{n}(y_i - \overline{y})(x_i - \overline{x})}{\sum_{i=1}^{n}(x_i - \overline{x})^2} \triangleq \frac{L_{xy}}{L_{xx}} \tag{6.3.6}$$

则直线$\hat{y} = \hat{\alpha} + \hat{\beta}x$就是我们要求的直线．由（6.3.5）式看出，点（$\overline{x}$，$\overline{y}$）在该直线上．

具体的计算可列表进行．对于例6.3.1有

$\sum x = 234$	$\sum y = 811.3$	$n = 9$
$\overline{x} = 26$	$\overline{y} = 90.1444$	
$(\sum x)^2/n = 6084$	$(\sum y)^2/n = 73134.188$	
$\sum x^2 = 10144$	$\sum y^2 = 76218.17$	
$\sum xy = 24628.6$	$(\sum x)(\sum y)/n = 21093.8$	

$$\hat{\beta} = \frac{3534.8}{4060} = 0.8706$$

$\hat{\alpha} = 67.5088$

$\hat{y} = 67.5088 + 0.8706x$

（二）数学模型与相关性检验

在上一段，我们利用n对数据（x_1，y_1），（x_2，y_2），……，（x_n，y_n）作出了一条回归直线$\hat{y} = \hat{\alpha} + \hat{\beta}x$。如果我们另作一组试验，又可取得$n$对数据，这样又可作出一条回归直线。一般说来，这两条回归直线是不相同的。因此，用这种办法得到的回归直线只能叫样本回归直线。这种叫法强调了它依赖于所取的子样，具有经验性质的特征。在实际生活中，我们所要考察的两个量往往都是随机变量。这时如何去研究它们之间的相关关系就不能不成为概率论与数理统计这一学科的研究对象了。为简单起见，我们在本章只讨论Y是随机变量，x是普通变量的情况，对于X，Y都是随机变量的情况，理论上需要新的处理方法，这里我们就不讨论了。但是，应该指出的是最后所得到的公式形式上却是一样的。因此从应用的角度，我们并不拘泥于要考虑X是哪一种类型的变量。

假定

$$Y = \alpha + \beta x + \varepsilon \tag{6.3.7}$$

这表明，自变量取值x时，因变量Y是随机变量$\alpha + \beta x + \varepsilon$，这里，$\alpha$，$\beta$是常数，$\varepsilon$是随机误差。对于随机误差$\varepsilon$，通常假定它服从正态$N(0, \sigma^2)$。对大多数情况是可以这样假设的。

为了得到α，β的估计，我们必须进行多次独立观测（或试验），设所得到的结果为

$$(x_1, Y_1), \quad (x_2, Y_2) \cdots, \quad (x_n, Y_n) \tag{6.3.8}$$

由（6.3.7）可知

$$Y_1 = \alpha + \beta x_1 + \varepsilon_1$$

$$Y_2 = \alpha + \beta x_2 + \varepsilon_2$$

$$\cdots$$

$$Y_n = \alpha + \beta x_n + \varepsilon_n$$

其中ε_1, ε_2, …, ε_n是相互独立的随机变量．它们均服从相同的正态分布$N(0, \sigma^2)$．这就是一元线性回归的数学模型．

对于一元线性回归模型，我们研究下述两个问题：

1　根据样本（6.3.8）去估计未知参数α, β；

2　对统计假设H_0: $\beta = 0$进行显著性检验．

由第一段可知，用最小二乘法，依据样本（6.3.8）我们可求出

$$\hat{\alpha} = \overline{Y} - \hat{\beta}\,\overline{x}$$

$$\hat{\beta} = \frac{L_{xY}}{L_{xx}}$$

式中，$\overline{Y}$, L_{xY}, $\hat{\alpha}$, $\hat{\beta}$均是随机变量，从而$\hat{Y}_i = \hat{\alpha} + \hat{\beta} x_i$也是随机变量，$1 \leqslant i \leqslant n$．我们称$\hat{\alpha}$, $\hat{\beta}$为未知参数α, β的最小二乘估计．

注意到

$$E(\hat{\beta}) = E\left(\frac{L_{xY}}{L_{xx}}\right) = \frac{1}{L_{xx}} E(L_{xY})$$

$$= \frac{1}{L_{xx}} E\left[\sum_{i=1}^{n} (x_i - \overline{x})(Y_i - \overline{Y})\right]$$

$$= \frac{1}{L_{xx}} \sum_{i=1}^{n} (x_i - \overline{x}) E(Y_i - \overline{Y})$$

但是，$EY_i = E(\alpha + \beta x_i + \varepsilon_i) = \alpha + \beta x_i$,

$$E\overline{Y} = E\left(\frac{1}{n}\sum_{i=1}^{n} Y_i\right) = \frac{1}{n}\sum_{i=1}^{n} EY_i = \alpha + \beta\,\overline{x}$$

所以

$$E\hat{\beta} = \frac{1}{L_{xx}} \sum_{i=1}^{n} (x_i - \overline{x})\beta(x_i - \overline{x}) = \beta$$

$$E\hat{\alpha} = E(\overline{Y} - \hat{\beta}\,\overline{x}) = E(\overline{Y}) - \overline{x}\,E\hat{\beta} = \alpha + \beta\overline{x} - \beta\overline{x} = \alpha$$

这表明$\hat{\alpha}$，$\hat{\beta}$分别是未知参数α，β的无偏估计.

下面，我们讨论关于假设H_0：$\beta=0$的显著性检验的问题. 这就是所谓的相关性检验. 由于在导出x与$\hat{y}$的定量关系式时，并不依赖于x与Y之间是否确实存在线性关系，只要有n对数据，代入公式（6.3.5），（6.3.6）中就可算出相应的$\hat{\alpha}$，$\hat{\beta}$来。显而易见，如果x与Y之间不存在线性关系，则作这样的演算就不过是数学游戏而已，毫无实际意义。为此，我们应该判明Y与x间是否确有线性相关关系. 在模型（6.3.7）下，就是要检验假设H_0：$\beta=0$。

为了给出合适的统计量，我们先导出偏差平方和的分解公式。

引理　对于任何n组数据(x_1, Y_1)，(x_2, Y_2)，…，(x_n, Y_n)恒成立

$$\sum_{i=1}^{n}(Y_i-\overline{Y})^2=\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2+\sum_{i=1}^{n}(\hat{Y}_i-\overline{Y})^2 \tag{6.3.9}$$

其中，$\hat{Y}_i=\hat{\alpha}+\hat{\beta}x_i$，　　$i=1,2,\cdots,n$

证明

$$\begin{aligned}\sum_{i=1}^{n}(Y_i-\overline{Y})^2&=\sum_{i=1}^{n}[(Y_i-\hat{Y}_i)+(\hat{Y}_i-Y)]^2\\&=\sum_{i=1}^{n}[Y_i-\hat{Y}_i]^2+\sum_{i=1}^{n}(\hat{Y}_i-\overline{Y})^2\\&\quad+2\sum_{i=1}^{n}(Y_i-\hat{Y}_i)(\hat{Y}_i-\overline{Y})\end{aligned}$$

但是，

$$\begin{aligned}&\sum_{i=1}^{n}(Y_i-\hat{Y}_i)(\hat{Y}_i-\overline{Y})\\&=\sum_{i=1}^{n}[Y_i-(\hat{\alpha}+\hat{\beta}x_i)][\hat{\alpha}+\hat{\beta}x_i-\overline{Y}]\\&=\sum_{i=1}^{n}[Y_i-\overline{Y}-\hat{\beta}(x_i-\overline{x})]\hat{\beta}(x_i-\overline{x})\end{aligned}$$

$$(\because \hat{\alpha} = \overline{Y} - \hat{\beta}\overline{x})$$

$$= \hat{\beta}\sum_{i=1}^{n}[(Y_i - \overline{Y})(x_i - \overline{x}) - \hat{\beta}(x_i - \overline{x})^2]$$

$$= \hat{\beta}\ (L_{xY} - \hat{\beta} L_{xx}) = 0$$

所以，（6.3.9）式成立。

对于这三个平方和，我们解释一下。

$\sum_{i=1}^{n}(Y_i - \overline{Y})^2$是$Y_1$，$Y_2$，…，$Y_n$ 这n个数据的偏差平方和，它反映了这n个数据的波动程度。记为L_{YY}。

$\hat{Y}_i = \hat{\alpha} + \hat{\beta}x_i$ 是回归直线$Y = \hat{\alpha} + \hat{\beta}x$上，横坐标为$x_i$的点的纵坐标，并且

$$\frac{1}{n}\sum_{i=1}^{n}\hat{Y}_i = \frac{1}{n}\sum_{i=1}^{n}(\hat{\alpha} + \hat{\beta}x_i) = \hat{\alpha} + \hat{\beta}\overline{x} = \overline{Y}$$

由此，平方和$\sum_{i=1}^{n}(\hat{Y}_i - \overline{Y})^2$就是$\hat{Y}_1$，$\hat{Y}_2$，…，$\hat{Y}_n$这$n$个数值的偏差平方和。它反映了$\hat{Y}_1$，$\hat{Y}_2$，…，$\hat{Y}_n$ 的波动程度，而这种波动主要来自Y与x之间的线性关系。我们称它为回归平方和，记为U。

平方和$\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2$在讲最小二乘原则时见到过。它是$Q(a, b)$的最小值，记作$Q$。它反映了扣除数据$Y$与$x$的线性关系外一切因素所引起的波动，称为残差平方和（或剩余平方和）。

由（6.3.9）式，$L_{YY} = U + Q$。

到此为止，我们把总的离差平方和分解成Q与U两部分，Q反映了由于随机误差所引起的观察值的波动，而U反映由于X与Y线性关系所引起的波动。因此，若检验β是否为0，只要用比值$\frac{U}{Q/(n-2)}$的大小来判断就行了。为此需要知道$\frac{U}{Q/(n-2)}$的分布。

若$\beta = 0$，则$Y = \alpha$，即所有的Y_i都服从正态分布$N(\alpha, \sigma^2)$，可以

证明，这时Q/σ^2，U/σ^2及L_{YY}/σ^2分别服从下列的χ^2分布，且Q与U相互独立。

Q/σ^2服从自由度为$(n-2)$的χ^2分布

U/σ^2服从自由度为1的χ^2分布

L_{YY}/σ^2服从自由度为$(n-1)$的χ^2分布

因此有下面结论，即

若$\beta=0$，则

$$F=\frac{U}{Q/(n-2)}$$

服从自由度为$(1, n-2)$的F分布。

如果$\beta\neq 0$，则U相对较大，这时F值也大，将F的观测值F_0与F分布的右侧临界值F_α进行比较，即可判断回归直线有无意义。

现在用上法对例6.3.1的相关性进行检验。

$$U=\sum_{i=1}^{n}(\hat{Y}_i-\overline{Y})^2=\sum_{i=1}^{n}[\hat{\alpha}+\hat{\beta}x_i-(\hat{\alpha}+\hat{\beta}\overline{x})]$$

$$=\sum_{i=1}^{n}\hat{\beta}^2(x_i-\overline{x})^2=\hat{\beta}^2L_{xx}=\hat{\beta}L_{xY}$$

$$Q=L_{YY}-U=\sum_{i=1}^{n}Y_i^2-n\overline{Y}^2-U$$

将具体数据代入计算得

$$F=\frac{U}{Q/(9-2)}=\frac{0.8706\times 3534.8}{0.9459}=3253.41$$

查F分布表，$\alpha=0.01$，自由度为1，7得$F_{0.01}=12.25$，现在$3253.41>12.25$，所以拒绝H_0：$\beta=0$，从而可认为x，Y有线性相关关系。

进行相关性的检验，也可用相关系数r，它定义为

$$r=\frac{L_{xY}}{\sqrt{L_{xx}\ L_{YY}}}$$

它反映了回归平方和在总平方和L_{YY}中的比例，显然$|r|$愈大，回

归效果愈好。由于U是L_{YY}中的一部分，而Q又非负，所以$U \leqslant L_{YY}$，因此可知$|r| \leqslant 1$。

当$r = 0$时，则$L_{YY} = 0$，因此$\beta = 0$说明x与Y无关。

当$0 < |r| < 1$时，则说明x与Y存在一定的线性关系。当$r > 0$时$\beta > 0$，这时x与Y正相关，〔图(6.3.2)〕，当$r < 0$时，$\beta < 0$，这时x与Y负相关〔图(6.3.3)〕，当$|r| = 1$时，$Q = L_{YY} - U = L_{YY}(1 - r^2) = 0$说明$n$个点$(x_i, Y_i)$全在回归直线上。这时$x$与$Y$之间存在确定的线性函数关系。

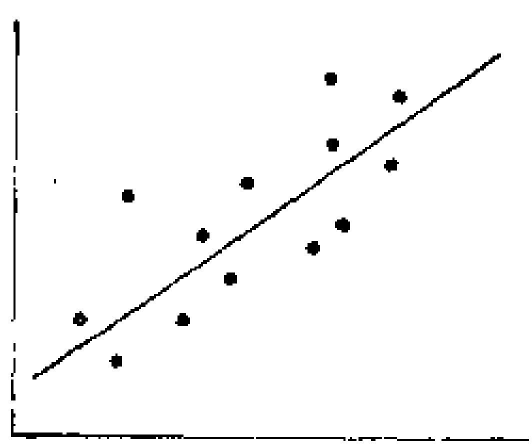


图 (6.3.2)

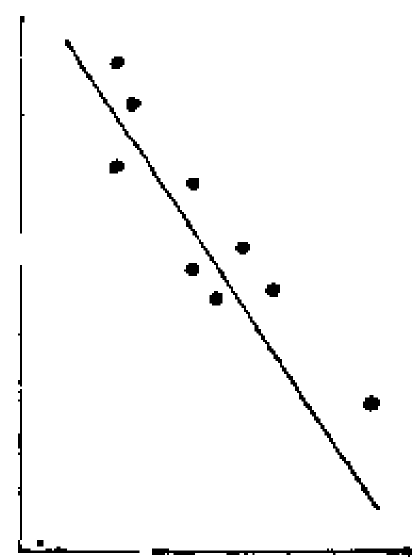


图 (6.3.3)

从上面分析可看出，相关系数r确实表示了x与Y之间的线性相关的密切程度。$|r|$愈接近于0，x与Y之间的线性相关程度愈小，反之，$|r|$愈大，愈接近于1，x与Y之间的线性相关就愈为密切。对一个具体问题，只有当$|r|$大到一定程度时才可认为x与Y有线性相关关系。书末的相关系数检验表给出了对不同的n，在两种显著水平α(0.05及0.01)下相关系数达到显著的最小值。求出的$|r|$值要大于表中的数，才能考虑用直线来描述x与Y之间的关系。例6.3.1中$n = 9$，查$n - 2 = 7$的一行，相应的数为0.798(1%)，而计算的

$$|r| = \sqrt{\frac{L_{xY}^2}{L_{xx}L_{YY}}} = \sqrt{\frac{U}{L_{YY}}} = 0.9989 > 0.798$$

所以同样得出结论 x 与 Y 有线性相关关系。

在对相关性进行检验时，可用 F 检验，也可用相关系数 r 进行检验，它们本质上是一样的。但相关系数的数值大小能够反映变量间线性相关的密切程度。

（三）根据回归方程进行预报

所谓预报，就是对任一给定的 x_0，推断相应的 y_0 值大致在一个什么范围里。对于任一给定的 $x=x_0$，由回归方程可以计算出一个回归值

$$\hat{Y}_0=\hat{\alpha}+\hat{\beta}x_0,$$

假定在 $x=x_0$ 时，Y 的观测值为 Y_0，则有

$$Y_0=\alpha+\beta x_0+\varepsilon_0$$

其中 $\varepsilon_0\sim N(0,\ \sigma^2)$，现在研究，对已给的置信度 $1-\alpha_0$，如何求一个正数 δ，使得 Y_0 以概率 $1-\alpha_0$ 落在区间 $(\hat{Y}_0-\delta,\ \hat{Y}_0+\delta)$ 里，即

$$P(\hat{Y}_0-\delta<Y_0<\hat{Y}_0+\delta)=1-\alpha_0$$

$$P(|Y_0-\hat{Y}_0|<\delta)=1-\alpha_0$$

为解决这个问题，应当求 $Y_0-\hat{Y}_0$ 的分布。因为 Y_0 与 $\hat{Y}_0$ 服从正态分布，所以 $Y_0-\hat{Y}_0$ 也服从正态分布，进而，我们可以算出。

$$E(Y_0-\hat{Y}_0)=0,$$

$$D(Y_0-\hat{Y}_0)=\sigma^2\left[1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{L_{xx}}\right]$$

所以有 $$Y_0-\hat{Y}_0\sim N\left(0,\sigma^2\left[1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{\Sigma(x_i-\bar{x})^2}\right]\right)$$

由于在分布中的 σ^2 未知，所以不能用正态分布做出 Y_0 的区间估计，这时，可用 $Q/(n-2)$ 代替 σ^2 而用统计量

$$\frac{(Y_0-\hat{Y}_0)^2}{\left[1+\frac{1}{n}+\frac{(x_0-\bar{x})^2}{\Sigma(x_i-\bar{x})^2}\right]\cdot\frac{Q}{n-2}}$$

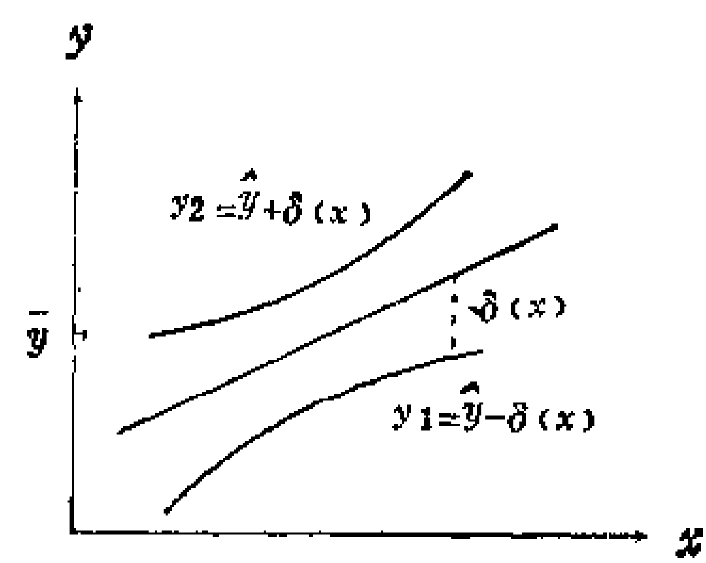
它服从自由度为（1，$n-2$）的F分布，由此不难求出Y_0的置信度为$1-\alpha_0$的置信区间为

$$(\hat{Y}_0-\delta,\ \hat{Y}_0+\delta)$$

而

$$\delta=\sqrt{F\alpha_0\frac{Q}{n-2}\left[1+\frac{1}{n}+\frac{(x_0-\overline{x})^2}{\Sigma(x_i-\overline{x})^2}\right]}$$

上式说明，Y的置信区间不仅与α_0有关，与n有关，而且与观测值x_0有关，当x_0靠近$\overline{x}$时，δ就小，因而区间就小，当x_0远离$\overline{x}$时，δ就大，因而区间就长，所以$\delta=\delta$（x_0）也是x_0的一个函数，$Y_1=\overline{Y}_0-\delta$和$Y_2=\overline{Y}_0+\delta$的图形如图（6.3.4）．



图（6.3.4）

下面将本节的主要内容概括在下表中。

1　进行相关性检验

（1）假设x，Y没有线性关系：$\beta=0$

（2）计算L_{xx}，L_{xY}，

$L_{xx}=\sum x_i^2-(\sum x_i)^2/n$

$L_{xY}=\sum x_iY_i-(\sum x_i)(\sum Y_i)/n$

（3）计算L_{YY}，Q，U

$L_{YY}=\Sigma(Y_i-\overline{Y})^2$

回归平方和$U=\hat{\beta}L_{XY}=L_{xY}^2/L_{xx}$

剩余平方和$Q=L_{YY}-U$

（4）求平方和的自由度

L_{YY}的自由度为$n-1$

U 的自由度为 1

Q 的自由度为$n-2$

（5）计算F_0值

$$F_0=\frac{U}{Q/(n-2)}$$

（6）判断

若 $F_0 \geqslant F_\alpha$，则认为$\beta \neq 0$

若 $F_0 < F_\alpha$ 则认为 $\beta=0$

若结论为$\beta \neq 0$，则按下面的步骤求回归直线

2 求回归直线

（1）求 a，b

$$b=L_{xY}/L_{xx} \qquad a=\overline{Y}-b\overline{X}$$

（2）回归直线为

$$Y=a+bx$$

3 计算相关系数

$$r=\frac{L_{xY}}{\sqrt{L_{xx}L_{YY}}}$$

4 预报

（1）对于给定的α，与已知自由度（1，$n-2$），求 F_α。

（2）对于给定的 x_0 求$\delta(x_0)$

$$\delta(x_0)=\sqrt{F_\alpha \cdot \frac{Q}{n-2}\left[1+\frac{1}{n}+\frac{(x_0-\overline{x})^2}{L_{xx}}\right]}$$

（3）置信度为 $1-\alpha$ 的置信区间为

$$(\widehat{Y}_0-\delta,\ \widehat{Y}_0+\delta)$$

*§6.4 多元线性回归

上一节讨论了两个变量的回归问题，其中因变量只与一个自变量有关。在很多实际问题中，影响因变量的因素不是一个而是多

个，这类回归问题称为多元回归问题。

例如，考察水稻的丰产性状。我们知道，水稻产量与它的三个产量构成因素，即每平方米的穗数，每穗实粒数，千粒重之间是密切相关的。找出它们之间的定量表达式无疑是有重大意义的。

又如，考察某个化学反应的速度。我们发现它与反应物的浓度 c 及所加的温度 t 有关。当 c 和 t 改变时，反应速度也随之改变。但是，由于种种偶然因素的影响，当 c 和 t 固定时，所观察到的反应速度也不完全相同，这时反应速度的观测值 Y 就是一个随机变量。最简单的情况是

$$Y=\beta_1 c+\beta_2 t+\varepsilon \tag{6.4.1}$$

其中 ε 是不可观测的随机变量。稍复杂一些的关系也可能是

$$Y=\beta_1 c+\beta_2 t+\beta_3 c^2+\beta_4 t^3+\beta_5 ct^2+\varepsilon \tag{6.4.2}$$

等等。值得注意的是，无论（6.4.1）式或是（6.4.2）式，Y 都是诸参数 β_i 的线性函数。只要我们通过试验或观测能够估计未知参数 β_1，β_2，…，则变量之间的定量关系式就找到了。

在这一节里，我们讨论多元线性回归问题。由于矩阵运算是多元的算术，所以应用矩阵工具是自然的。我们可以看到，利用矩阵工具，很多结果均可以用简洁的形式表出，并且证明也不繁杂。一元线性回归是多元线性回归的特例。我们在这一节将顺便给出上一节的全部理论结果及其证明。

（一）多元线性回归的数学模型

设自变量为 $C_1,C_2,\cdots,C_p$，因变量为 Y，n 次观测数据为 $(c_{i1}, c_{i2},\cdots,c_{ip},Y_i)$ $i=1,2,\cdots,n$。假定

$$Y_i=\beta_1 c_{i1}+\beta_2 c_{i2}+\cdots+\beta_p c_{ip}+\varepsilon_i \tag{6.4.3}$$

式中，β_1，β_2，…，β_p 为未知参数，ε_1，ε_2，…，ε_n 是 n 个相互独立的随机变量，服从相同的正态分布 $N(0,\sigma^2)$。

利用矩阵、向量的记法，令

$$
Y=\begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}_{n\times 1}, \qquad \beta=\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{pmatrix}_{p\times 1}
$$

$$
C=\begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ & \cdots\cdots & & \\ c_{n1} & c_{n2} & \cdots & c_{np} \end{pmatrix}_{n\times p} \qquad \varepsilon=\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \cdots \\ \varepsilon_n \end{pmatrix}_{n\times 1}
$$

有

$$
\begin{cases} Y=C\beta+\varepsilon \\ \varepsilon\sim N(0,\ \sigma^2 I_n) \end{cases} \tag{6.4.4}
$$

注意到，

$$
E(Y)=E(C\beta+\varepsilon)=C\beta+E(\varepsilon)=C\beta
$$

$$
D(Y)=E(Y-C\beta)(Y-C\beta)'=E\varepsilon\varepsilon'=\sigma^2 I_n
$$

（6.4.4）式又可写为

$$
\begin{cases} E(Y)=C\beta \\ D(Y)=\sigma^2 I_n \end{cases} \tag{6.4.5}
$$

上述（6.4.3），（6.4.4），（6.4.5）这三种写法，表达了同一件事情．以后，我们称其为多元线性回归的数学模型，记为（Y，$C\beta$、$\sigma^2 I_n$）．

我们要考虑的统计推断问题是：

1　对未知参数向量 β 及未知参数σ^2进行估计；

2　检验关于参数向量 β 的某些假设；

3　对Y作预测．

本节的所有讨论均假定$n>p$，且$rk(C'C)=p$.这时矩阵 C列满秩，$(C'C)^{-1}$存在．当 $C'C$是奇异矩阵的情况，需要广义逆矩阵等新的代数工具，在这里我们就不讨论了．有兴趣的读者可参看有关的专著．

（二）β的最小二乘估计及其性质

设β_1，β_2，…，β_p的估计分别记为$\hat{\beta}_1$，$\hat{\beta}_2$，…，$\hat{\beta}_p$，称 p 元线性方程

$$\hat{Y}=\hat{\beta}_1C_1+\hat{\beta}_2C_2+\cdots+\hat{\beta}_pC_p$$

为 p 元线性回归方程。对每一组自变量的值（$c_{i1},c_{i2},\cdots,c_{ip}$），由上式可求得回归值

$$\hat{y}_i=\hat{\beta}_1c_{i1}+\hat{\beta}_2c_{i2}+\cdots+\hat{\beta}_pc_{ip},\qquad i=1,\ 2,\cdots,\ n$$

我们希望由估计$\hat{\beta}_1$，$\hat{\beta}_2$，…，$\hat{\beta}_p$定出的回归方程能使所有的y_i与$\hat{y}_i$之间的偏差平方和为最小，即要求

$$\min_{\beta_1;\beta_2,\cdots\beta_p}\sum_{i=1}^{n}(y_i-\beta_1c_{i1}-\beta_2c_{i2}-\cdots-\beta_pc_{ip})^2$$

$$=\sum_{i=1}^{n}(y_i-\hat{\beta}_1c_{i1}-\hat{\beta}_2c_{i2}-\cdots-\hat{\beta}_pc_{ip})^2$$

定义未知参数β_1，β_2，…，β_p的 p 元函数

$$Q(\beta_1,\ \beta_2,\ \cdots,\ \beta_p)=\sum_{i=1}^{n}(y_i-\beta_1c_{i1}-\beta_2c_{i2}-\cdots$$
$$-\beta_pc_{ip})^2.$$

只要我们能求出使$Q(\beta_1,\ \beta_2,\ \cdots,\ \beta_p)$达到极小值的$\beta_1=\hat{\beta}_1$，$\beta_2=\hat{\beta}_2$，…，$\beta_p=\hat{\beta}_p$问题就解决了。用微积分学中求极值的理论，令

$$\frac{\partial Q(\beta_1,\ \beta_2,\ \cdots,\ \beta_p)}{\partial\beta_j}=0\qquad j=1,2,\cdots,p$$

求解这一方程组即得。

注意到

$$\frac{\partial Q(\beta_1,\ \beta_2,\ \cdots,\ \beta_p)}{\partial\beta_j}=-2\sum_{i=1}^{n}(y_i-\beta_1c_{i1}-\cdots-\beta_pc_{ip})c_{ij},$$
$$j=1,2,\cdots,p,$$

因此，要求的$\hat{\beta}_j$应满足线性方程组

$$\sum_{i=1}^{n} c_{ij}y_i=\left(\sum_{i=1}^{n} c_{i1}c_{ij}\right)\beta_1+\left(\sum_{i=1}^{n} c_{i2}c_{ij}\right)\beta_2+\cdots+$$

$$\left(\sum_{i=1}^{n} c_{ip}c_{ij}\right)\beta_p \qquad j=1,2,\cdots,p$$

但上式左端恰为向量$C'Y$的第j个元素，而$\left(\sum_{i=1}^{n} c_{i1}c_{ij},\right.$

$\left.\sum_{i=1}^{n} c_{i2}c_{ij}, \cdots, \sum_{i=1}^{n} c_{ip}c_{ij}\right)$恰为矩阵$C'C$的第$j$行。因此，写成矩阵形式为

$$C'C\hat{\beta}=C'Y \tag{6.4.6}$$

称为正规方程组。我们已经假定系数阵$C'C$是满秩的。所以

$$\hat{\beta}=(C'C)^{-1}C'Y \tag{6.4.7}$$

应当注意的是，用最小二乘法，导出正规方程组，求得最小二乘估计$\hat{\beta}$，虽然是对一组确定的数据（y_i，c_{i1}，c_{i2}，$\cdots$，c_{ip}）作的，但是在模型(6.4.5)下，y_i是可以波动的，即Y_i为随机变量。由此（6.4.7）中的Y是n维随机向量，从而$\hat{\beta}$是p维随机向量。

下面我们讨论$\hat{\beta}$的性质。

性质 1 $\quad E(\hat{\beta})=\beta$

证明 $\quad E(\hat{\beta})=E[(C'C)^{-1}C'Y]$

$$=(C'C)^{-1}C'EY=(C'C)^{-1}C'C\beta=\beta$$

这表明β的最小二乘估计$\hat{\beta}$是β的无偏估计。性质1用$\hat{\beta}$的分量表示就是

$$E(\hat{\beta}_i)=\beta_i,\ i=1,2,\cdots,p$$

性质 2 $\quad D(\hat{\beta})=\sigma^2(C'C)^{-1}$

证明 $\quad D(\hat{\beta})=D[(C'C)^{-1}C'Y]$

$$=(C'C)^{-1}C'\cdot DY\cdot C(C'C)^{-1}$$

$$=\sigma^2(C'C)^{-1}C'I_nC(C'C)^{-1}$$

$$=\sigma^2(C'C)^{-1}$$

性质 3　$E(Y-C\hat{\beta})=0$

$cov(Y-C\hat{\beta},\ \hat{\beta})=0$。

证明　$E(Y-C\hat{\beta})=EY-E(C\hat{\beta})=C\beta-CE\hat{\beta}=0$

$$\begin{aligned}
&cov(Y-C\hat{\beta},\ \hat{\beta})\\
&=cov((I-C(C'C)^{-1}C')Y,\ (C'C)^{-1}C'Y)\\
&=(I-C(C'C)^{-1}C')DY\cdot C(C'C)^{-1}\\
&=\sigma^2(I-C(C'C)^{-1}C')C(C'C)^{-1}\\
&=0
\end{aligned}$$

向量 $Y-C\hat{\beta}=Y-\hat{Y}$ 称为剩余向量（或称残差向量）．性质 3 表明剩余向量的均值是 0 向量，并且它与向量 $\hat{\beta}$ 不相关．

记　$S_e^2=(Y-\hat{Y})'(Y-\hat{Y})=\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2$，称为剩余平方和（或残差平方和）．

残差平方和 S_e^2 还可写成各种不同的形式：

$$\begin{aligned}
S_e^2&=(Y-C\hat{\beta})'(Y-C\hat{\beta})\\
&=(Y-C(C'C)^{-1}C'Y)'(Y-C(C'C)^{-1}C'Y)\\
&=Y'[I-C(C'C)^{-1}C'][I-C(C'C)^{-1}C']Y\\
&=Y'[I-C(C'C)^{-1}C']Y\\
&=Y'Y-Y'C(C'C)^{-1}C'Y\\
&=Y'Y-Y'C(C'C)^{-1}C'C(C'C)^{-1}C'Y\\
&=Y'Y-\hat{\beta}'C'C\hat{\beta}
\end{aligned}$$

性质 4　$E(S_e^2)=(n-p)\sigma^2$

证明　$$\begin{aligned}
E(S_e^2)&=E((Y-\hat{Y})'(Y-\hat{Y}))\\
&=E[tr(Y-\hat{Y})'(Y-\hat{Y})]\\
&=E[tr(Y-\hat{Y})(Y-\hat{Y})']\\
&=trE[(Y-\hat{Y})(Y-\hat{Y})']
\end{aligned}$$

由性质 3 知，$E(Y-\hat{Y})=0$，故

$$E(Y-\hat{Y})(Y-\hat{Y})'=D(Y-\hat{Y})=D[Y-C(C'C)^{-1}C'Y]$$

$$=[I-C(C'C)^{-1}C']DY[I-C(C'C)^{-1}C']$$
$$=\sigma^2[I-C(C'C)^{-1}C']$$

由此，

$$\begin{aligned}E(S_e^2)&=\sigma^2 tr[I_n-C(C'C)^{-1}C']\\&=\sigma^2[trI_n-trC(C'C)^{-1}C']\\&=\sigma^2[n-tr(C'C)^{-1}C'C]\\&=\sigma^2[n-trI_p]\\&=\sigma^2(n-p)\end{aligned}$$

记

$$\hat{\sigma}_e^2=\frac{S_e^2}{n-p}$$

由性质4有

$$E(\hat{\sigma}_e^2)=E\left(\frac{S_e^2}{n-p}\right)=\frac{1}{n-p}E(S_e^2)=\sigma^2$$

即$\hat{\sigma}_e^2$是σ^2的无偏估计．以后，我们称$\hat{\sigma}_e^2$为剩余方差．

细心的读者可能已经发现，在上述四条性质的证明中，均未用到关于ε所服从的分布的假定，只是应用了关于它的一阶矩与二阶矩的假定，$E(\varepsilon)=0, D(\varepsilon)=\sigma^2 I_n$．下面，我们假定$\varepsilon\sim N(0,\sigma^2 I_n)$，证明一条十分重要的定理．

定理6.4.1　当$Y\sim N(C\beta,\ \sigma^2 I_n)$时，$\hat{\beta}$与$S_e^2$相互独立，并且

（1）　$\hat{\beta}\sim N(\beta,\ \sigma^2(C'C)^{-1})$

（2）　$Y-C\hat{\beta}\sim N(0,\ \sigma^2(I_n-C(C'C)^{-1}C'))$

（3）　$S_e^2/\sigma^2\sim\chi^2(n-p)$

证明

（1）　$\hat{\beta}=(C'C)^{-1}C'Y$ 是相互独立的n个正态随机变量Y_1，Y_2，…，Y_n的线性函数，因此，$\hat{\beta}$服从正态分布．性质1．性质2已经给出

$$E(\hat{\beta})=\beta$$
$$D(\hat{\beta})=\sigma^2(C'C)^{-1}$$

所以，

$$\hat{\beta} \sim N(\beta,\ \sigma^2(C'C)^{-1})$$

（2） $Y-C\hat{\beta}=Y-C(C'C)^{-1}C'Y=[I_n-C(C'C)^{-1}C']Y$ 是服从正态分布的 n 维随机向量 Y 的线性函数，所以 $Y-C\hat{\beta}$ 服从正态分布。性质 3，性质 4 已经给出

$$E(Y-C\hat{\beta})=0,$$

$$D(Y-C\hat{\beta})=\sigma^2[I_n-C(C'C)^{-1}C']$$

所以

$$Y-C\hat{\beta} \sim N(0,\ \sigma^2(I_n-C(C'C)^{-1}C'))$$

由性质 3 知，$\hat{\beta}$ 与 $Y-C\hat{\beta}$ 是不相关的，在正态分布的场合，不相关与相互独立是等价的，故 $\hat{\beta}$ 与 $Y-C\hat{\beta}$ 相互独立，但 S_e^2 仅是 $Y-C\hat{\beta}$ 的函数，所以 $\hat{\beta}$ 与 S_e^2 相互独立。

（3） $S_e^2=Y'[I_n-C(C'C)^{-1}C']Y$

要证明 S_e^2/σ^2 服从自由度为 $n-p$ 的 χ^2 分布，只要设法将 S_e^2 表为 $n-p$ 个相互独立的 $N(0,\ \sigma^2)$ 变量的平方和即可。

令 $P_C=C(C'C)^{-1}C'$，直接计算知 $P_C^2=P_C$，即 P_C 是幂等阵，且 P_C 还是对称阵。由

$$rkC=rk(C(C'C)^{-1}C'C)=rkP_CC \leqslant rkP_C=rkC(C'C)^{-1}C' \leqslant rkC$$

知 $rkP_C=rkC=p$。因为对称幂等阵一定是非负定的，而幂等阵的特征根非 0 即 1，所以存在正交阵 Γ，使得

$$\Gamma P_C\Gamma'=\begin{pmatrix} I_p & 0 \\ 0 & 0 \end{pmatrix}$$

作线性变换 $Z=\Gamma(Y-C\beta)=\Gamma\varepsilon$，由于 Z 是正态变量 ε 的线性函数，所以 Z 仍服从正态分布。

注意到，

$$E(Z)=E(\Gamma\varepsilon)=\Gamma E(\varepsilon)=0$$

$$D(Z)=D(\Gamma\varepsilon)=\Gamma(D\varepsilon)\Gamma'=\Gamma(\sigma^2I_n)\Gamma'=\sigma^2\Gamma\Gamma'=\sigma^2I_n$$

这表明 $Z\sim N(0,\ \sigma^2I_n)$，即 Z 的 n 个分量 $Z_1,\ Z_2,\ \cdots,\ Z_n$ 相互独

立，且服从相同的正态分布$N(0,\sigma^2)$。

由$\boldsymbol{Z}=\Gamma(\boldsymbol{Y}-C\boldsymbol{\beta})$与$\Gamma'\Gamma=\boldsymbol{I}_n$推得$\boldsymbol{Y}=\Gamma'\boldsymbol{Z}+C\boldsymbol{\beta}$，于是

$$
\begin{aligned}
S_e^2&=\boldsymbol{Y}'[\boldsymbol{I}_n-\boldsymbol{P}_C]\boldsymbol{Y}\\
&=(\Gamma'\boldsymbol{Z}+C\boldsymbol{\beta})'[\boldsymbol{I}_n-\boldsymbol{P}_C](\Gamma'\boldsymbol{Z}+C\boldsymbol{\beta})
\end{aligned}
$$

注意到$(\boldsymbol{I}_n-\boldsymbol{P}_C)C=C-\boldsymbol{P}_CC=C-C(C'C)^{-1}C'C=0$，所以

$$
\begin{aligned}
S_e^2&=\boldsymbol{Z}'\Gamma\Gamma'\boldsymbol{Z}-\boldsymbol{Z}'\Gamma\boldsymbol{P}_C\Gamma\boldsymbol{Z}\\
&=\boldsymbol{Z}'\boldsymbol{I}_n\boldsymbol{Z}-\boldsymbol{Z}'\begin{pmatrix}\boldsymbol{I}_p & 0\\ 0 & 0\end{pmatrix}\boldsymbol{Z}\\
&=Z_1^2+Z_2^2+\cdots+Z_n^2-(Z_1^2+Z_2^2+\cdots+Z_p^2)\\
&=Z_{p+1}^2+Z_{p+2}^2+\cdots+Z_n^2
\end{aligned}
$$

这是$n-p$个相互独立的，正态 $N(0,\sigma^2)$ 的随机变量的平方和，所以$S_e^2/\sigma^2\sim\chi^2(n-p)$。

作为这一段有关理论结果的一个应用，我们回到一元线性回归的情形，顺便得到上一段中遗留下来的未加证明的一些结果的证明。

一元线性回归的数学模型是

$$
\begin{cases}
Y_1=\alpha+\beta x_1+\varepsilon_1\\
Y_2=\alpha+\beta x_2+\varepsilon_2\\
\qquad\cdots\cdots\\
Y_n=\alpha+\beta x_n+\varepsilon_n
\end{cases}
$$

写成矩阵的形式为

$$
\begin{cases}
E\boldsymbol{Y}=C\boldsymbol{\beta}\\
D\boldsymbol{Y}=\sigma^2\boldsymbol{I}_n,
\end{cases}
$$

其中，

$$
\boldsymbol{Y}=\begin{pmatrix}Y_1\\Y_2\\\vdots\\Y_n\end{pmatrix},\quad C=\begin{pmatrix}1 & x_1\\1 & x_2\\\vdots & \vdots\\1 & x_n\end{pmatrix},\quad \boldsymbol{\beta}=\begin{pmatrix}\alpha\\\beta\end{pmatrix},\quad \boldsymbol{\varepsilon}=\begin{pmatrix}\varepsilon_1\\\varepsilon_2\\\vdots\\\varepsilon_n\end{pmatrix}
$$

由此

$$C'C=\begin{pmatrix}1 & 1 & \cdots & 1\\ x_1 & x_2 & \cdots & x_n\end{pmatrix}\begin{pmatrix}1 & x_1\\ 1 & x_2\\ \vdots & \vdots\\ 1 & x_n\end{pmatrix}=\begin{pmatrix}n & n\overline{x}\\ n\overline{x} & \sum_{i=1}^{n}x_i^2\end{pmatrix},$$

$$|C'C|=n\sum_{i=1}^{n}x_i^2-n^2\overline{x}^2=n\left(\sum_{i=1}^{n}x_i^2-n\overline{x}^2\right)$$

$$=n\sum_{i=1}^{n}(x_i-\overline{x})^2$$

$$(C'C)^{-1}=\frac{1}{|C'C|}\begin{pmatrix}\sum_{i=1}^{n}x_i^2 & -n\overline{x}\\ -n\overline{x} & n\end{pmatrix},$$

$$=\frac{1}{n\sum_{i=1}^{n}(x_i-\overline{x})^2}\begin{pmatrix}\sum_{i=1}^{n}x_i^2 & -n\overline{x}\\ -n\overline{x} & n\end{pmatrix}$$

$$\hat{\beta}=(C'C)^{-1}C'Y=\frac{1}{n\sum_{i=1}^{n}(x_i-\overline{x})^2}$$

$$\begin{pmatrix}\sum_{i=1}^{n}x_i^2 & -n\overline{x}\\ -n\overline{x} & n\end{pmatrix}\begin{pmatrix}1 & 1 & \cdots & 1\\ x_1 & x_2 & \cdots & x_n\end{pmatrix}\begin{pmatrix}Y_1\\ Y_2\\ \vdots\\ Y_n\end{pmatrix}$$

$$=\frac{1}{n\sum_{i=1}^{n}(x_i-\overline{x})^2}\begin{pmatrix}\sum_{i=1}^{n}x_i^2 & -n\overline{x}\\ -n\overline{x} & n\end{pmatrix}\begin{pmatrix}n\overline{Y}\\ \sum_{i=1}^{n}x_iY_i\end{pmatrix}$$

$$=\frac{1}{n\sum_{i=1}^{n}(x_i-\overline{x})^2}\begin{pmatrix}n\overline{Y}\sum_{i=1}^{n}x_i^2-n\overline{x}\sum_{i=1}^{n}x_iY_i\\ n\sum_{i=1}^{n}x_iY_i-n^2\overline{x}\,\overline{Y}\end{pmatrix}$$

由此得

$$\hat{\beta} = \frac{1}{n\sum_{i=1}^{n}(x_i - \bar{x})^2}\left(n\sum_{i=1}^{n}x_iY_i - n^2\,\bar{x}\,\bar{Y}\right)$$

$$= \frac{\sum_{i=1}^{n}x_iY_i - n\,\bar{x}\,\bar{Y}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$= \frac{\sum_{i=1}^{n}(Y_i - \bar{Y})(x_i - \bar{x})}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

$$\hat{\alpha} = \frac{1}{n\sum_{i=1}^{n}(x_i - \bar{x})^2}\left(n\bar{Y}\sum_{i=1}^{n}x_i^2 - n\,\bar{x}\sum_{i=1}^{n}x_iY_i\right)$$

$$= \frac{\bar{Y}\left(\sum_{i=1}^{n}x_i^2 - n\bar{x}^2\right) - \bar{x}\left(\sum_{i=1}^{n}x_iY_i - n\,\bar{x}\,\bar{Y}\right)}{\sum_{i=1}^{n}x_i^2 - n\bar{x}^2}$$

$$= \bar{Y} - \hat{\beta}\,\bar{x}$$

由定理6.4.1知，$\hat{\beta} \sim N(\beta,\ \sigma^2(C'C)^{-1})$，表成分量的形式，即

$$\hat{\alpha} \sim N\left(\alpha,\ \sigma^2\frac{\sum_{i=1}^{n}x_i^2}{n\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)$$

$$\hat{\beta} \sim N\left(\beta,\ \sigma^2\frac{1}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right)$$

$$cov(\hat{\alpha},\ \hat{\beta}) = -\frac{\sigma^2\bar{x}}{\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

除此而外，我们还知道，

$$\frac{S_e^2}{\sigma^2}=\frac{\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2}{\sigma^2}\sim\chi^2(n-2)$$

由于

$$\sum_{i=1}^{n}(Y_i-\alpha)^2=\sum_{i=1}^{n}[(Y_i-\overline{Y})+(\overline{Y}-\alpha)]^2$$

$$=\sum_{i=1}^{n}(Y_i-\overline{Y})^2+n(\overline{Y}-\alpha)^2$$

$$=\sum_{i=1}^{n}(Y_i-\hat{Y}_i)^2+\sum_{i=1}^{n}(\hat{Y}_i-\overline{Y})^2$$

$$+n(\overline{Y}-\alpha)^2,$$

注意到，在假设H_0：$\beta=0$成立时，诸$Y_i\sim N(\alpha,\ \sigma^2)$，$i=1,2,\cdots,n$，所以

$$\frac{\sum_{i=1}^{n}(Y_i-\alpha)^2}{\sigma^2}\sim\chi^2(n),$$

$$\frac{(\overline{Y}-\alpha)^2}{\frac{\sigma^2}{n}}\sim\chi^2(1).$$

因为

$$U=\sum_{i=1}^{n}(\hat{Y}_i-\overline{Y})^2=\sum_{i=1}^{n}[\hat{\alpha}+\hat{\beta}x_i-(\hat{\alpha}+\hat{\beta}\overline{x})]^2$$

$$=\sum_{i=1}^{n}\hat{\beta}^2(x_i-\overline{x})^2=\hat{\beta}^2\sum_{i=1}^{n}(x_i-\overline{x})^2=\hat{\beta}^2L_{xx}$$

但是，$\hat{\beta}\sim N\left(\beta,\ \frac{\sigma^2}{L_{xx}}\right)$，即$(\hat{\beta}-\beta)\Big/\frac{\sigma}{\sqrt{L_{xx}}}\sim N(0,\ 1)$

从而

$$\left(\frac{\hat{\beta}-\beta}{\frac{\sigma}{\sqrt{L_{xx}}}}\right)^2 \sim \chi^2(1)$$

当H_0：$\beta=0$成立时，$\hat{\beta}^2 L_{xx}/\sigma^2 \sim \chi^2(1)$．此外，已经知道$S_e^2/\sigma^2 \sim \chi^2(n-2)$，注意到各个平方和均是非负定的二次型，且它们的秩之间成立等式

$$n=(n-2)+1+1$$

所以，由柯赫伦定理知，U和S_e^2相互独立，根据第三章定理3.5.11知统计量

$$F=\frac{U/1}{S_e^2/(n-2)} \sim F(1,\ n-2)$$

在上一节就是利用这一F统计量去作关于假设H_0：$\beta=0$的显著性检验的。

在考虑利用一元线性回归方程作统计予报时，关键在于寻求$Y_0-\hat{Y}_0$的分布。

由于Y_0与Y_1，…，Y_n相互独立，注意到$\hat{Y}_0=(1\quad x_0)\begin{pmatrix}\hat{\alpha}\\\hat{\beta}\end{pmatrix}=(1\quad x_0)\hat{\boldsymbol{\beta}}=(1\quad x_0)(C'C)^{-1}C'Y$是$Y$的线性函数，所以$Y_0$与$\hat{Y}_0$相互独立，且$\hat{Y}_0$也服从正态分布．$Y_0-\hat{Y}_0$是相互独立的两个正态变量的差，服从正态分布，它的两个参数为

$$\begin{aligned}E(Y_0-\hat{Y}_0)&=EY_0-E\hat{Y}_0=\alpha+\beta x_0-E(\hat{\alpha}+\hat{\beta}x_0)\\&=\alpha+\beta x_0-E\hat{\alpha}-x_0E\hat{\beta}\\&=\alpha+\beta x_0-\alpha-x_0\beta=0\end{aligned}$$

$$\begin{aligned}D(Y_0-\hat{Y}_0)&=DY_0+D\hat{Y}_0=\sigma^2+D(\hat{\alpha}+\hat{\beta}x_0)\\&=\sigma^2+D(\hat{\alpha})+x_0^2D\hat{\beta}+2x_0cov(\hat{\alpha},\ \hat{\beta})\\&=\sigma^2+\sigma^2\frac{\sum_{i=1}^n x_i^2}{n\sum_{i=1}^n(x_i-\bar{x})^2}\end{aligned}$$

$$+\sigma^2 \frac{x_0^2}{\sum_{i=1}^{n}(x_i-\overline{x})^2} - \sigma^2 \frac{2x_0\overline{x}}{\sum_{i=1}^{n}(x_i-\overline{x})^2}$$

$$=\sigma^2+\sigma^2\frac{\sum_{i=1}^{n}(x_i-x_0)^2}{nL_{xx}}$$

$$=\sigma^2\left[1+\frac{1}{n}+\frac{(x_0-\overline{x})^2}{L_{xx}}\right]$$

由此，

$$Y_0-\hat{Y}_0\sim N\left(0,\ \sigma^2\left(1+\frac{1}{n}+\frac{(x_0-\overline{x})^2}{L_{xx}}\right)\right)$$

当 σ^2 未知时，我们用它的无偏估计量 $S_\varepsilon^2/(n-2)$ 代之，由于 $Y_0-\hat{Y}_0$ 与 S_ε^2 相互独立，所以，

$$\frac{\dfrac{(Y_0-\hat{Y}_0)^2}{\sigma^2\left(1+\dfrac{1}{n}+\dfrac{(x_0-\overline{x})^2}{L_{xx}}\right)}\Big/1}{\dfrac{S_\varepsilon^2}{\sigma^2}/(n-2)}\sim F(1,\ n-2)$$

这样，要求出 Y_0 的置信度为 $1-\alpha$ 的置信区间就不困难了。具体的作法上一节已经讲过，这里不再重复。

（三）关于参数 β 的假设检验

在讨论一元线性回归时，我们已经讲过，因变量与自变量之间存在线性相关关系只是一种假设。如果这一假设不成立，求线性回归方程是毫无意义的。因此，我们必须对线性关系的假定进行检验。类似的问题在多元的情形也存在。我们需要检验随机变量 Y 与一组自变量 C_1，C_2，…，C_k 间存在着线性关系的假设。具体的作法是，提出假设

$$H_0: \beta_1 = \beta_2 = \cdots = \beta_p = 0$$

如果H_0不否定，则表明Y与C_1，C_2，…，C_p间不存在线性关系。如何去导出相应于这一检验问题的统计量呢？我们仍然从平方和的分解公式入手。

$$Y'Y = (Y - C\hat{\beta} + C\hat{\beta})'(Y - C\hat{\beta} + C\hat{\beta})$$
$$= (Y - C\hat{\beta})'(Y - C\hat{\beta}) + \hat{\beta}'C'C\hat{\beta}$$

交叉项等于0是因为，

$$(Y - C\hat{\beta})'C\hat{\beta} = \hat{\beta}'C'(Y - C\hat{\beta})$$
$$= \hat{\beta}'C'[Y - C(C'C)^{-1}C'Y] = \hat{\beta}'C'[I_n - C(C'C)^{-1}C']Y$$
$$= \hat{\beta}'[C' - C'C(C'C)^{-1}C']Y = 0$$

记$S_e^2 = (Y - C\hat{\beta})'(Y - C\hat{\beta})$，$U = \hat{\beta}'C'C\hat{\beta}$，分别称为误差平方和与回归平方和．于是，

$$Y'Y = S_e^2 + U$$

注意到，这三个平方和皆是Y_1，Y_2，…，Y_n的非负定二次型。

$Y'Y$的秩为$rkI_n = n$，

$S_e^2 = (Y - C\hat{\beta})'(Y - C\hat{\beta}) = Y'[I_n - C(C'C)^{-1}C']Y$，它的秩为$rk[I_n - C(C'C)^{-1}C'] = n - p$，

$U = \hat{\beta}'C'C\hat{\beta} = Y'C(C'C)^{-1}C'C(C'C)^{-1}C'Y$
$= Y'C(C'C)^{-1}C'Y$，它的秩为$rk[C(C'C)^{-1}C'] = tr[C(C'C)^{-1}C']$
$= tr[(C'C)^{-1}C'C] = p$。

当假设H_0：$\beta = 0$成立时，$Y \sim N(0, \sigma^2 I_n)$，这时，

$$\frac{Y'Y}{\sigma^2} = \frac{1}{\sigma^2}\sum_{i=1}^{n} Y_i^2 \sim \chi^2(n)$$

$$\frac{S_e^2}{\sigma^2} \sim \chi^2(n-p)$$

$$\frac{U}{\sigma^2} \sim \chi^2(p)$$

且S_e^2与U相互独立．因而统计量

$$F = \frac{U/p}{S_e^2/(n-p)} \sim F(p,\ n-p)$$

检验统计量的分布求出后，就可以按通常程序进行显著性检验了，这里不再赘述。

如果经过检验后认为Y与C_1，C_2，…，C_p存在线性关系，进一步的问题是哪些自变量更为重要。换句话说，有的自变量的系数β很大，而有的自变量β的系数很小甚至等于0．因此，我们还应对假设H_0：$\beta_i=0$进行显著性检验。

我们知道，$\hat{\beta} \sim N(\beta,\ \sigma^2(C'C)^{-1})$，并且$\hat{\beta}$与$S_e^2$相互独立。记逆矩阵$(C'C)^{-1}$的第$(i,\ j)$元素为$c^{ij}$，则有$\hat{\beta}$的第$i$个分量

$$\hat{\beta}_i \sim N(\beta_i,\ \sigma^2 c^{ii})。$$

由此，当假设H_0成立时，则有

$$\frac{\dfrac{\hat{\beta}_i^2}{c^{ii}}}{\dfrac{S_e^2}{n-p}} \sim F(1,\ n-p)$$

往下，显著性检验的程序与其它检验问题是一样的，这里不再赘述。

经过这样的显著性检验后，我们就可以从众多的自变量中筛选出较重要的自变量，再在这些自变量与Y间建立回归方程。这样得到的回归方程比较简洁，更加有利于揭示所要研究的对象的本质。

（四）对Y的预测

所谓预测问题，指的是知道了自变量C_1，C_2，…，C_p的取值状态，要估计Y的取值范围．在一元的情况，我们已经讨论过了，这里讨论多元情形下的预测．处理这个问题的总的思路和以前的办法是一样的。

给定$C_0'=(c_{01},\ c_{02},\ \cdots,\ c_{0p})$的值，要求$Y_0$的区间估计．由于$Y_0$与$Y_1$，$Y_2$，…，$Y_n$相互独立，$\hat{Y}_0$与$S_e^2$相互独立，$Y_0$与$\hat{Y}_0$相互独立，$Y_0$与$S_e^2$也相互独立，因此$Y_0-\hat{Y}_0$与$S_e^2$相互独立。

已经知道，$Y_0 \sim N(C'_0\beta,\ \sigma^2)$，$\hat{Y}_0 \sim N(C'_0\beta, \sigma^2 C'_0(C'C)^{-1}C_0)$，所以$Y_0-\hat{Y}_0$也服从正态分布，且

$$E(Y_0-\hat{Y}_0)=EY_0-E\hat{Y}_0=C'_0\beta-C'_0\beta=0$$

$$\begin{aligned}D(Y_0-\hat{Y}_0)&=D(Y_0)+D(\hat{Y}_0)=\sigma^2+D(C'_0\hat{\beta})\\&=\sigma^2+C'_0D(\hat{\beta})C_0\\&=\sigma^2[1+C'_0(C'C)^{-1}C_0]\end{aligned}$$

但是，

$$\frac{S_e^2}{\sigma^2}\sim\chi^2(n-p)$$

且$Y_0-\hat{Y}_0$与S_e^2相互独立，所以

$$T=\frac{\dfrac{Y_0-\hat{Y}_0}{\sigma\sqrt{1+C'_0(C'C)^{-1}C_0}}}{\dfrac{S_e}{\sigma\sqrt{n-p}}}\sim t(n-p)$$

给定显著性水平α，查自由度为$n-p$的t分布表得临界值 t_α，满足

$$P(|t|\geqslant t_\alpha)=\alpha$$

由此

$$P\left(|Y_0-\hat{Y}_0|\leqslant t_\alpha\cdot\frac{S_e}{\sqrt{n-p}}\cdot\sqrt{1+C'_0(C'C)^{-1}C_0}\right)$$

$$=1-\alpha$$

令 $\delta=t_\alpha\dfrac{S_e}{\sqrt{n-p}}\cdot\sqrt{1+C'_0(C'C)^{-1}C_0}$ 则

$$P(\hat{Y}_0-\delta\leqslant Y_0\leqslant\hat{Y}_0+\delta)=1-\alpha$$

这个式子表明Y_0的置信水平为$1-\alpha$的预测区域是由两个超平面间所围成的区域。

（五）一个实例

平炉炼钢过程中，由于矿石及炉气的氧化作用，铁水的总含碳量不断降低。一炉钢在冶炼初期（熔化期）中总的去碳量Y与所加的两种矿石（天然矿石与烧结矿石）的加入量x_1，x_2及熔化时间x_3

（熔化时间愈长则去碳量愈多）有关。经实测某号平炉的相应数据如下表，求Y对x_1，x_2，x_3的线性回归方程，并作统计分析。

编号	Y（吨）	x_1（槽）	x_2（槽）	x_3（5分钟）	编号	Y（吨）	x_1（槽）	x_2（槽）	x_3（5分钟）
1	4.3302	2	18	50	26	2.7066	9	6	39
2	3.6485	7	9	40	27	5.6314	12	5	51
3	4.4830	5	14	46	28	5.8152	6	13	41
4	5.5468	12	3	43	29	5.1302	12	7	47
5	5.4970	1	20	64	30	5.3910	0	24	61
6	3.1125	3	12	40	31	4.4583	5	12	37
7	5.1182	3	17	64	32	4.6569	4	15	49
8	3.8759	6	5	39	33	4.5212	0	20	45
9	4.6700	7	8	37	34	4.8650	6	16	42
10	4.9536	0	23	55	35	5.3566	4	17	48
11	5.0060	3	16	60	36	4.6098	10	4	48
12	5.2701	0	18	49	37	2.3815	4	14	36
13	5.3772	8	4	50	38	3.8746	5	13	36
14	5.4849	6	14	51	39	4.5919	9	8	51
15	4.5960	0	21	51	40	5.1588	6	13	54
16	5.6645	3	14	51	41	5.4373	5	8	100
17	6.0795	7	12	56	42	3.9960	5	11	44
18	3.2194	16	0	48	43	4.3970	8	6	63
19	5.8076	6	16	45	44	4.0622	2	13	55
20	4.7306	0	15	52	45	2.2905	7	8	50
21	4.6805	9	0	40	46	4.7115	4	10	45
22	3.1272	4	6	32	47	4.5310	10	5	40
23	2.6104	0	17	47	48	5.3637	3	17	64
24	3.7174	9	0	44	49	6.0771	4	15	72
25	3.8946	2	16	39					

令

$$\hat{Y}=\hat{\beta}_0+\hat{\beta}_1x_1+\hat{\beta}_2x_2+\hat{\beta}_3x_3.$$

由于存在着常数项β_0，在具体计算时，可将上式改写为

$$\hat{Y}_i=\hat{\mu}+\hat{\beta}_1(x_{i1}-\overline{x_1})+\hat{\beta}_2(x_{i2}-\overline{x}_2)+\hat{\beta}_3(x_{i3}-\overline{x}_3)$$

其中，

$$\overline{x}_j\triangleq\frac{1}{n}\sum_{i=1}^{n}x_{ij},\qquad j=1,2,3$$

$$\hat{\mu}\triangleq\hat{\beta}_0+\hat{\beta}_1\overline{x}_1+\hat{\beta}_2\overline{x}_2+\hat{\beta}_3\overline{x}_3$$

确定$\hat{\mu}$，$\hat{\beta}_1$，$\hat{\beta}_2$，$\hat{\beta}_3$的正规方程组为

$$\begin{pmatrix} n & 0 & 0 & 0 \\ 0 & L_{11} & L_{12} & L_{13} \\ 0 & L_{21} & L_{22} & L_{23} \\ 0 & L_{31} & L_{32} & L_{33} \end{pmatrix}\begin{pmatrix} \hat{\mu} \\ \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix}=\begin{pmatrix} n\overline{Y} \\ L_{01} \\ L_{02} \\ L_{03} \end{pmatrix}$$

其中，

$$\overline{Y}\triangleq\frac{1}{n}\sum_{i=1}^{n}Y_i$$

$$L_{0j}\triangleq\sum_{i=1}^{n}(Y_i-\overline{Y})(x_{ij}-\overline{x}_j),\quad j=1,2,3$$

$$L_{jk}\triangleq\sum_{i=1}^{n}(x_{ij}-\overline{x}_j)(x_{ik}-\overline{x}_k),\qquad j,k=1,2,3.$$

由此，正规方程又可写为

$$\hat{\mu}=\overline{Y}$$

$$\begin{pmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{pmatrix}\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix}=\begin{pmatrix} L_{01} \\ L_{02} \\ L_{03} \end{pmatrix}$$

记 $L_{00}=\sum_{i=1}^{n}(Y_i-\overline{Y})^2$，则可算出

$$S_e^2 = \sum_{i=1}^{n} [Y_i - \hat{Y}_i]^2$$

$$= \sum_{i=1}^{n} [Y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 x_{i2} + \hat{\beta}_3 x_{i3})]^2$$

$$= \sum_{i=1}^{n} [(Y_i - \overline{Y}) - \sum_{j=1}^{3} \hat{\beta}_j (x_{ij} - \overline{x}_j)]^2$$

由于，

$$\sum_{j=1}^{3} \hat{\beta}_j (x_{ij} - \overline{x}_j)$$

$$= (x_{i1} - \overline{x}_1, x_{i2} - \overline{x}_2, x_{i3} - \overline{x}_3) \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix}$$

$$= (\hat{\beta}_1, \quad \hat{\beta}_2, \quad \hat{\beta}_3) \begin{pmatrix} x_{i1} - \overline{x}_1 \\ x_{i2} - \overline{x}_2 \\ x_{i3} - \overline{x}_3 \end{pmatrix}$$

所以，

$$\sum_{i=1}^{n} \left[\sum_{j=1}^{3} \hat{\beta}_j (x_{ij} - \overline{x}_j) \right]^2$$

$$= \sum_{i=1}^{n} (\hat{\beta}_1, \quad \hat{\beta}_2, \quad \hat{\beta}_3) \begin{pmatrix} x_{i1} - \overline{x}_1 \\ x_{i2} - \overline{x}_2 \\ x_{i3} - \overline{x}_3 \end{pmatrix}$$

$$(x_{i1} - \overline{x}_1, \quad x_{i2} - \overline{x}_2, \quad x_{i3} - \overline{x}_3) \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix}$$

$$
=(\hat{\beta}_1,\ \hat{\beta}_2,\ \hat{\beta}_3)\begin{pmatrix} \sum_{i=1}^{n}(x_{i1}-\overline{x}_1)^2 & \sum_{i=1}^{n}(x_{i1}-\overline{x}_1)(x_{i2}-\overline{x}_2) & \sum_{i=1}^{n}(x_{i1}-\overline{x}_1)(x_{i3}-\overline{x}_3) \\ \sum_{i=1}^{n}(x_{i2}-\overline{x}_2)(x_{i1}-\overline{x}_1) & \sum_{i=1}^{n}(x_{i2}-\overline{x}_2)^2 & \sum_{i=1}^{n}(x_{i2}-\overline{x}_2)(x_{i3}-\overline{x}_3) \\ \sum_{i=1}^{n}(x_{i3}-\overline{x}_3)(x_{i1}-\overline{x}_1) & \sum_{i=1}^{n}(x_{i3}-\overline{x}_3)(x_{i2}-\overline{x}_2) & \sum_{i=1}^{n}(x_{i3}-\overline{x}_3)^2 \end{pmatrix}\begin{pmatrix}\hat{\beta}_1\\ \hat{\beta}_2\\ \hat{\beta}_3\end{pmatrix}
$$

$$
=(\hat{\beta}_1,\ \hat{\beta}_2,\ \hat{\beta}_3)\begin{pmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{pmatrix}\begin{pmatrix}\hat{\beta}_1\\ \hat{\beta}_2\\ \hat{\beta}_3\end{pmatrix}
$$

$$
=(\hat{\beta}_1,\ \hat{\beta}_2,\ \hat{\beta}_3)\begin{pmatrix} L_{01}\\ L_{02}\\ L_{03}\end{pmatrix}=\sum_{j=1}^{3}\hat{\beta}_j L_{0j}
$$

这样一来，

$$
\begin{aligned}
S_e^2 &= \sum_{i=1}^{n}(Y_i-\overline{Y})^2 \\
&- 2\sum_{j=1}^{3}\hat{\beta}_j\left[\sum_{i=1}^{n}(Y_i-\overline{Y})(x_{ij}-\overline{x}_j)\right]+\sum_{j=1}^{3}\hat{\beta}_j L_{0j} \\
&= \sum_{i=1}^{n}(Y_i-\overline{Y})^2-2\sum_{j=1}^{3}\hat{\beta}_j L_{0j}+\sum_{j=1}^{3}\hat{\beta}_j L_{0j} \\
&= L_{00}-\sum_{j=1}^{3}\hat{\beta}_j L_{0j}
\end{aligned}
$$

由表中的数据计算得

$$n=49,\quad \sum Y_i=224.5169,\quad \overline{Y}=4.582$$

$$\sum x_{i1}=259,\quad \overline{x}_1=5.286$$

$$\sum x_{i2}=578,\quad \overline{x}_2=11.796$$

$$\sum x_{i3}=2411,\quad \overline{x}_3=49.204$$

$$L_{11}=662.00,\quad L_{12}=-918.1428,\quad L_{13}=-388.8571$$

$$L_{22}=1753.959,\quad L_{23}=776.0408,\quad L_{33}=6247.959$$

$$L_{01}=-6.4329,\quad L_{02}=69.1305,\quad L_{03}=245.5713$$

代入正规方程得

$$\begin{pmatrix} 662.00 & -918.1428 & -388.8571 \\ -918.1428 & 1753.959 & 776.0408 \\ -388.8571 & 776.0408 & 6247.959 \end{pmatrix}\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix}=\begin{pmatrix} -6.4329 \\ 69.1305 \\ 245.5713 \end{pmatrix}$$

$$\hat{\beta}_0=4.582-5.286\beta_1-11.796\beta_2-49.204\beta_3$$

解之得

$$\hat{\beta}_1=0.1606,\quad \hat{\beta}_2=0.1076,$$

$$\hat{\beta}_3=0.0359,\quad \hat{\beta}_0=0.6974.$$

于是总去碳量Y对两种矿石加入量 x_1，x_2及熔化时间x_3的线性回归方程为

$$\hat{Y}=0.6974+0.1606x_1+0.1076x_2+0.0359x_3$$

再由

$$S_e^2=L_{00}-\sum_{i=1}^{3}\hat{\beta}_i L_{0i}=44.905-15.230=29.675$$

得

$$\hat{\sigma}^2=\frac{S_e^2}{n-p}=\frac{29.675}{49-4}=0.659$$

$$\hat{\sigma}=0.81$$

下面我们考虑关于未知参数的检验问题。首先检验假设 H_0: $\beta_1=\beta_2=\beta_3=0$。这时所用的统计量为

$$F=\frac{U/3}{S_e^2/(n-p)}\sim F(3,45)$$

其中 $U=\sum_{j=1}^{3}\hat{\beta}_j L_{0j}$。将具体数据代入进行计算

$$u=15.230,\quad \hat{\sigma}^2=0.659$$

于是 F 的数值为 7.7036。在显著水平 $\alpha=0.01$ 下，查 F 分布表得临界值 $F_{0.01}(3,45)=4.24$，由于 $7.7036>4.24$，故拒绝 H_0，即去碳量 Y 与 x_1，x_2，x_3 的回归方程在 $\alpha=0.01$ 下有显著意义。

下面再对各个回归系数进行显著性检验。为此，首先算出逆矩阵

$$\begin{pmatrix} L_{11} & L_{12} & L_{13} \\ L_{21} & L_{22} & L_{23} \\ L_{31} & L_{32} & L_{33} \end{pmatrix}^{-1}$$

的对角元素。它们分别是

$$c^{11}=0.005515,\quad c^{22}=0.002122,\quad c^{33}=0.0001694$$

检验假设 H_{0j}：$\beta_j=0$，$j=1,2,3$ 的 F 统计量的数值分别为

$$\frac{\hat{\beta}_1^2}{c^{11}\hat{\sigma}^2}=7.1289,\quad \frac{\hat{\beta}_2^2}{c^{22}\hat{\sigma}^2}=8.2944,\quad \frac{\hat{\beta}_3^2}{c^{33}\hat{\sigma}^2}=11.6281$$

取 $F=0.05$，查 F 分布表得 $F_{0.05}(1,45)=4.05$。由于上述数值全大于4.05，所以对 $j=1,2,3$ 均拒绝 H_{0j}，即认为三个变量对 Y 均有显著影响。

习　　题

1　某农业科学试验站进行了一项玉米品种对比试验，试验结

果为（单位：斤/小区）

产量 品种 重复	Ⅰ	Ⅱ	Ⅲ	Ⅳ
1	32.3	33.3	30.8	29.3
2	34.0	33.0	34.3	26.0
3	34.3	36.3	35.3	29.8
4	35.0	36.8	32.3	28.0
5	36.5	34.5	35.8	28.8

问这四个不同的玉米品种的平均产量间是否有显著差异？$(\alpha=0.01)$

2　抽查某地区三所小学五年级男学生的身高，得数据如下：

小　学	身		高	数		据
一	128.1	134.1	133.1	138.9	140.8	127.4
二	150.3	147.9	136.8	126.0	150.7	155.8
三	140.6	143.1	144.5	143.7	148.5	146.4

问该地区这三所小学五年级男学生的平均身高是否有显著差别？$(\alpha=0.05)$

3　下表记录了三位操作工分别在四台不同的机器上操作三天的日产量：

机器	操作工 1	操作工 2	操作工 3
M_1	15，15，17	19，19，16	16，18，21
M_2	17，17，17	15，15，15	19，22，22
M_3	15，17，16	18，17，16	18，18，18
M_4	18，20，22	15，16，17	17，17，17

试检验：

（1）操作工之间的差异是否显著？

（2）机器之间的差异是否显著？

（3）交互作用是否显著？（$\alpha=0.05$）

4　某水稻研究所进行了一项水稻栽培试验，考察株行距与育苗方式对丰产性能的影响。试验数据如下：（单位：斤/小区）

产量 株行距 苗别	B_1	B_2	B_3	B_4
	7 × 3	8 × 3	9 × 3	10 × 3
A_1 薄膜苗	45.4，48.0	47.0，45.0	43.0，45.2	47.6，49.2
A_2 湿润苗	44.1，40.1	41.7，37.5	36.9，40.9	36.0，40.2

试检验：

（1）不同株行距之间的差异是否显著？

（2）不同育苗方式之间的差异是否显著？

（3）交互作用是否显著？（$\alpha=0.01$）

5　某医院用光电比色计检验尿汞时，得尿汞含量（mg/l）与消光系数读数的结果如下：

尿汞含量（x）	2	4	6	8	10
消光系数（y）	64	138	205	285	360

已知它们之间服从线性模型

$$EY_i=\alpha+\beta x_i$$

试求α和β的最小二乘估计，并检验β是否为0？（$\alpha=0.01$）

6　某种合成纤维的强度与其拉伸倍数有线性关系，实测24个纤维样品的强度与相应的拉伸倍数数据如下：

拉伸倍数x	1.9	2.0	2.1	2.5	2.7	2.7	3.5	3.5	4.0	4.0	4.5	4.6
强　度y	1.4	1.3	1.8	2.5	2.8	2.5	3.0	2.7	4.0	3.5	4.2	3.5

拉伸倍数 x	5.0	5.2	6.0	6.3	6.5	7.1	8.0	8.0	8.9	9.0	9.5	10.0
强　度 y	5.5	5.0	5.5	6.4	6.0	5.3	6.5	7.0	8.5	8.0	8.1	8.1

求强度对拉伸倍数的回归直线方程，并计算样本相关系数，进行相关性检验。$(\alpha=0.01)$

7　某农业科学试验站为了了解某个小麦新品种的增产情况，在一定的种植条件下，研究千粒重与产量的相关关系，数据如下

千粒重（X）	33.9	30.3	32.5	29.4	27.1	34.8	31.2	34.6	32.9
产　量（Y）	3182.5	2461.1	3318	1973	1436	3230.5	3130	3583.5	3624.7

求样本相关系数，并进行相关性检验。$(\alpha=0.05)$

8　某大队为了提高蔬菜生产的产量，用1:1000的青虫菌药液防治菜青虫。由于一次喷药并不能将菜青虫全部消灭，残留的菜青虫还会继续繁殖，需要找到喷药后菜青虫个数(Y)与喷药后天数(X)之间的关系以确定第二次喷药何时为宜。取160棵甘蓝作样本，每天数一次菜青虫的个数得下表：

喷药后天数	0	1	2	3	4	5	6	7	8	9	10	11
菜青虫头数	23	28	33	40	48	57	68	81	98	116	142	167

试确定Y与X之间的关系。（提示：考虑 $Y=AB^X$，系数 A，B 待定）。

9　设

$$Y_1=a+\varepsilon_1$$

$$Y_2=2a-b+\varepsilon_2$$

$$Y_3=a+2b+\varepsilon_3$$

其中 $\varepsilon_i(i=1,2,3)$ 相互独立，且 $E(\varepsilon_i)=0$，$D(\varepsilon_i)=\sigma^2$。

（1）将上述关系式写成矩阵形式 $Y=C\beta+\varepsilon$，写出系数阵 C 与向量 β；

（2）求 $C'C$；

（3）写出正规方程；

（4）求出最小二乘估计 $\hat{\beta}$；

（5）求出 $D(\hat{a}),D(\hat{b})$。

10　今有四个物体，按下述方法称量，得到如下数据：

x_1	x_2	x_3	x_4	Y
1	1	1	1	20.2
1	-1	1	-1	8.0
1	1	-1	-1	9.2
1	-1	-1	1	1.4

其中 1 表示该物体放在天平左端，-1 表示该物体放在天平的右端，Y 是使天平达到平衡时，在右端所加砝码的重量，用最小二乘法估计这四个物体的重量 β_i，$i=1,2,3,4$。

学 习 指 导

（一）本章概要

方差分析与回归分析是数理统计学中最常用的两种统计分析方法．考虑到读者的不同兴趣与要求，在正文中，理论与方法是分开写的．对于从事统计方法应用的同志完全可以跳过带星号的章节不看，并不影响学会应用这些方法．

本章最基本的内容是单因子方差分析与一元回归分析．只有把它们真正学懂了，才有可能去学习多因子方差分析与多元回归分

析．读者应该首先学好这两部分的内容，结合实例，搞清它的直观想法，尔后通过练习，学会应用这些方法．在此基础上，有可能时，再去深究它的理论基础，这才是可行的途径．

从本质上讲，方差分析与回归分析都是属于线性模型统计分析的内容的．对于有较好的线性代数与矩阵论基础的读者完全可以采用类似于第四节的处理办法，把本章的内容总结与概括起来，并为进一步的学习与研究打下良好的基础．

（二）内容说明

1 关于农业科学试验的几个问题

方差分析最初是从处理农业科学试验的数据中提出来的一种统计分析方法．现在它已在工业、农业、科学实验的各个领域中得到了广泛的应用．

为了提供一些实际的背景材料，我们讲一下农业科学试验的几个有关的问题．

第一个问题，试验为什么要设置重复？

例如，试验的目的是考察杀螟杆菌和青虫菌防治菜青虫的效果．在田间划出两个地块进行对比试验，结果是，用杀螟杆菌，菜青虫的死亡率为89.5％，而用青虫菌的，为84.2％．如果你由此就下结论，杀螟杆菌比青虫菌效果好，显然太性急了．一次试验的结果，会受到种种偶然因素的干扰，因而往往会出假象．我们多作几次试验，每种药各用在四个地块里（即各重复4次），结果是杀螟杆菌，82.3％，青虫菌85.0％，得到与前相反的数据．以后我们多次试验的结果均表明，青虫菌药效确实高于杀螟杆菌．

这一试验的结果清楚地表明，对于科学试验必须有非常严肃的科学态度，单凭一次试验结果往往是不能作出正确结论的，必须经过多次试验之后才能作出结论．能否认为这个地方作一次试验，那个地方作一次试验，合起来就是多次试验了呢？这要看试验的要求与问题的性质而定．实际中的许多问题都带有因地制宜的性质．照

搬别人的经验并不总是行得通的。这样，试验的重复，主要是指的在当时当地的重复。设置重复的主要目的是减少偶然因素的作用，以更好地反映事物的必然规律。

既然试验设置重复是减少试验误差的重要措施之一，是否重复次数越多越好？当然，在一定条件下，重复次数越多，试验结果越可靠；但是，重复太多，人力物力消耗大，工作中失误的可能性增加，统计、调查的工作量加大，实际上反而达不到减少误差的目的。因此，一般的农业试验，均要求重复3—5次。从数理统计学的角度看，进行方差分析时，误差平方和S_e^2的自由度一般要求大于10，由此可确定出相应的抽样次数或重复次数。

第二个问题，为什么要设置区组？

农业田间试验既要有重复，又要求除试验因子外其它试验条件完全一致，这就带来了一个很大的矛盾。地力的不匀如何解决？重复次数越多，占地面积越大，地力越不均匀。解决这一矛盾的办法是进行局部控制，设置区组。我们把试验地分成若干个大块，使得每一个大块的地力基本一致，这样的大块就称为区组。如果我们要考察n个品种的丰产性状，我们就可以把每一区组分成n个小块（称为小区），尔后一个小区种一个品种，这种作法使得不同品种间的比较都是在大体上相同的条件下进行的，因而可以大大地减少试验误差。

第三个问题，为什么要采用随机排列？

进行品种对比试验时，各个品种在每一区组中的排列方式，通常有两种作法：一是顺序排列，一是随机排列。

由于试验不可避免地存在许多偶然因素的干扰，因而只有采用随机排列才能保证每一品种受到干扰的可能性是一样的。例如作物会受到风的影响，如果顺序排列，由西向东排甲、乙、丙，…，这样在经常刮西风的地区，总是甲品种最先迎风，乙品种次之，…，这样各个品种迎风的情况就不均衡了。如果采用随机排列，各个品种间的顺序是随机的，这个问题就解决了。

应该强调指出，不能把“随机”误解为“随意”．“随意”是主观的设想，带着主观的框框去作试验，偏爱某一种结果是不可能得出正确的结论来的。

随机排列的方法可以用抽签，查随机数表等．只要进行随机试验，那么它的各个等可能的结果就可以用来“翻译”成因子的各个水平。

当区组的大小（即每一区组中小区的个数）等于水平数时就称为完全的随机区组试验，如果区组的大小小于水平数时，就称为不完全的随机区组试验．后一种情况的试验设计问题引出了数学上许多有趣的难题。

2　关于符号$\overline{Y}_{i\cdot}$，$\overline{Y}_{i\cdot\cdot}$，…等

这一章里使用的符号大部分在统计文献上是通用的．k 个数据 Y_{i1}，Y_{i2}，…，Y_{ik} 的平均记为

$$\overline{Y}_{i\cdot}=\frac{1}{k}\sum_{j=1}^{k}Y_{ij}.$$

这一记法便于记忆．在Y的上方加“—”表示平均．对哪些数求算术平均呢？看脚标上“·”所处的位置．如果“·”写在记 j 的位置就是对变动 j 时的所有数据作平均．按照这种写法，

$$\frac{1}{mnl}\sum_{i=1}^{m}\sum_{j=1}^{n}\sum_{k=1}^{l}Y_{ijk}$$

应该写成$\overline{Y}_{\cdots}$，的确是这样，不少书就是这样写的，但我们已经习惯了用$\overline{Y}$来表示所有数据的总平均，并且略去“·”又不至于引起混淆，何乐而不为呢？

3 关于交互作用的概念

为了解释交互作用的概念，我们讲一个具体例子。

甲、乙两个工程队各自挖土方，甲队每天挖20方，乙队每天挖30方．把这两个队调到一起开挖土方，是否每天挖20＋30＝50（方）呢？实际的情况并不这么简单．可能挖60方，也可能只挖45方．前

一种情况是这两个队互相促进，提高了劳动效率；后一种情况是这两个队互相扯皮，降低了劳动效率。这表明，两队之间的相互关系（称为交互作用）也会影响挖方的进度。换言之，如果把每天的挖土方数(例如60)拆开，除了甲、乙两队各自的作用(=20+30)外，还有一个新的因素，即甲、乙两队间的交互作用（=10）。

一般地讲，在任何一个多因子试验中，除了各个因子的作用之外，还存在每两个因子之间的交互作用、每三个因子之间的交互作用等等。这样一来，在数据模型中无外乎增加了不少未知参数。因此要想考虑交互作用就得多作试验，并且还要多设重复。

4　关于回归分析

一元回归又称直线回归，在实际生活中有着广泛的应用。不仅变量Y与变量x之间存在线性相关关系时可以用，而且对相当一类曲线回归，只要作一适当的数据变换也可以化为直线回归加以解决。

在计算方法课程中学过的曲线拟合问题，在作法上类似于这里的回归分析，但那里处理的是两组确定的数据（因变量并不看成随机变量），没有涉及误差的分析，也不进行回归系数的显著性检验。当试验数据波动时，如何处置所得出的不同的样本回归曲线；利用配出的回归直线（经验公式）进行予测时，精度如何估计，所有这些，都只能在数理统计学的范畴中加以解决。

为什么要叫回归分析？

关于“回归”（regression）一词的使用要追溯到1889年生物学家格尔顿（F·Galton）与1903年统计学家皮尔逊（K·Pearson）的工作。他们在遗传学的研究中发现，后代的属性与双亲是相关的，但平均地说，在程度上要差一点。说得具体一点，皮尔逊收集了上千个家庭关于儿子的身高（Y）与父亲身高（X）的资料，作出的回归直线为$\hat{Y}=0.516X+33.73$。这项成果表明，虽然高个的父亲确有生高个儿子的趋势，但父亲身高增加一个单位时，儿子身高仅增加半个单位左右。平均说来，一群高个父亲的儿子们的平均高度要

低于他们的父亲的平均高度，而一群低个父亲的儿子们的高度要高于他们的父亲的平均高度。这就是说，高个子的儿子们的高度有一个回归，即回归到全种族男子平均身高的趋势。这种规律保证了一个种族的种种特征的存在。如果不是这样，高个子的父亲生出更高个子的儿子，而高个的儿子又生出更高的孙子，……如此继续下去，一个种族的平均身高就将失去意义了。

在遗传学的研究中，这样的例子是屡见不鲜的。某农业科学研究所研究水稻的遗传规律，得出杂交一代抽穗日数（从出芽到抽穗的总天数）Y 与它们双亲的平均抽穗日数 X 的回归方程为

$$Y = 19.08 + 0.8X$$

这一规律表明，双亲平均抽穗日数长的后代的抽穗日数平均说来要比它们的双亲短一点，而双亲平均抽穗日数短的后代平均说来要比它们的双亲长一点，有向它们的平均抽穗日数回归的趋势。

*5　关于方阵迹的一些性质

一个方阵 $A=(a_{ij})_{n\times n}$ 的迹 trA 定义为它的主对角线上元素的和，即

$$trA = \sum_{i=1}^{n} a_{ii}$$

特别，当 A 是一行一列矩阵，即是一个数 a 时有

$$tra = a$$

这一性质在正文的证明中经常用到。例如，由于二次型 $Y'AY$ 是一个数，所以

$$Y'AY = trY'AY$$

迹有下述性质：

（1）$tr(A+B) = trA + trB$

证明　若 $A=(a_{ij})_{n\times n}$，$B=(b_{ij})_{n\times n}$，则

$$A+B=(a_{ij}+b_{ij})_{n\times n},$$

因此，

$$tr\,(A+B) = \sum_{i=1}^{n}(a_{ii}+b_{ii}) = \sum_{i=1}^{n} a_{ii} + \sum_{i=1}^{n} b_{ii}$$

$$= trA + trB$$

（2）$tr(\lambda A) = \lambda trA$，$\lambda$是一个常数。

证明　$\lambda A = (\lambda a_{ij})_{n\times n}$，所以

$$tr\ (\lambda A) = \sum_{i=1}^{n} \lambda a_{ii} = \lambda \sum_{i=1}^{n} a_{ii} = \lambda trA$$

（3）$trA\cdot B = trB\cdot A$。

证明　设$A = (a_{ij})_{m\times n}$，$B = (b_{jk})_{n\times m}$，则

$$AB = \left(\sum_{j=1}^{n} a_{ij} b_{jk}\right)_{m\times m},\quad BA = \left(\sum_{k=1}^{m} b_{jk} a_{ki}\right)_{n\times n}$$

因此，

$$trAB = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} b_{ji},\quad trBA = \sum_{j=1}^{n}\sum_{i=1}^{m} b_{ji} a_{ij}$$

但是

$$\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij} b_{ji} = \sum_{j=1}^{n}\sum_{i=1}^{m} b_{ji} a_{ij}$$

所以

$$trA\cdot B = trB\cdot A$$

（4）$E(trX) = tr(EX)$

证明　若$X = (X_{ij})_{n\times n}$，则 $E(trX) = E\left(\sum_{i=1}^{n} X_{ii}\right)$，由于$n$个随机变量和的均值等于各个随机变量均值的和，即 $E\left(\sum_{i=1}^{n} X_{ii}\right) = \sum_{i=1}^{n} EX_{ii} = tr(EX)$。

*6　关于幂等阵的性质

一个方阵 $A = (a_{ij})_{n\times n}$，如果满足 $A^2 = A\cdot A = A$，则称为幂等阵。例如$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$就是一个幂等阵。

幂等阵具有下述性质：

（1）幂等阵的特征根非0即1。

证明　若k是幂等阵A的特征根，则有

$$A\alpha = k\alpha$$

其中α是非0向量，它是A的属于k的特征向量。两端同乘矩阵A得

$$A^2\alpha = kA\alpha$$

$$\Rightarrow \quad A\alpha = k\cdot k\alpha \quad (\because A^2 = A)$$

$$\Rightarrow \quad k\alpha = k^2\alpha$$

$$\Rightarrow \quad k(k-1)\alpha = 0$$

$$\Rightarrow \quad k = 0 \text{或} k = 1 \quad (\because \alpha \neq 0)$$

（2）对称幂等阵一定是非负定阵。

证明　若$A' = A$，且$A^2 = A$，则对任意向量x有

$$x'Ax = x'A^2x = x'A\cdot Ax = x'A'\cdot Ax$$

$$= (Ax)'\cdot Ax = \| Ax \|^2 \geqslant 0$$

这里$\|\cdot\|$表示（列）向量的欧氏范数，即若$y' = (y_1, y_2, \cdots, y_n)$，

$\| y \|^2 = y'y = \sum_{i=1}^{n} y_i^2$。由非负定阵的定义知$A$非负定。

（3）若矩阵A对称、幂等，则$rkA = trA$。

证明　设$rkA = r$，由于A是对称阵，故存在正交阵Γ，使得

$$A = \Gamma \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \Gamma'$$

其中$\lambda_1, \lambda_2, \cdots, \lambda_n$是$A$的特征根。注意到$A$又是秩为$r$的幂等阵，而幂等阵的特征根非0即1，所以$\lambda_1, \lambda_2, \cdots, \lambda_n$中恰有$r$个1，其余$n-r$个全为0。于是

$$trA = tr\left\{ \Gamma \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \Gamma' \right\} = tr\left\{ \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix} \Gamma'\Gamma \right\}$$

$$= tr\begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{pmatrix}$$

$$= \sum_{i=1}^{n} \lambda_i$$

$$= r = rkA$$

习 题 解 答

1 设四个品种的小区产量Y_1，Y_2，Y_3，Y_4为相互独立，且有相同方差的正态$N(\mu_i,\ \sigma^2)$ $i=1,2,3,4$,变量。

要求检验假设H_0：$\mu_1=\mu_2=\mu_3=\mu_4$。

将题给数据均减去30列入下表进行计算

	Ⅰ	Ⅱ	Ⅲ	Ⅳ	左边各数相加
	2.3	3.3	0.8	−0.7	
	4.0	3.0	4.3	−4.0	
	4.3	6.3	5.3	−0.2	
	5.0	6.8	2.3	−2	
	6.5	4.5	5.8	−1.2	
$\sum_j y_{ij}$	22.1	23.9	18.5	−8.1	56.4
$(\sum_j y_{ij})^2$	488.41	571.21	342.25	65.61	1467.48
$\sum_j y_{ij}^2$	107.03	126.07	86.15	21.97	341.22

由上表可得：

$$S_A^2 = \frac{1}{5}\sum_{i=1}^{4}\left(\sum_{j=1}^{5} y_{ij}\right)^2 - \frac{1}{5\times 4}\left(\sum_{i=1}^{4}\sum_{j=1}^{5} y_{ij}\right)^2$$

$$= \frac{1}{5}\times 1467.48 - \frac{1}{20}(56.4)^2 = 134.448$$

$$S_e^2 = \sum_{i=1}^{4}\sum_{j=1}^{5} y_{ij}^2 - \frac{1}{5}\sum_{i=1}^{4}\left(\sum_{j=1}^{5} y_{ij}\right)^2$$

$$= 341.22 - \frac{1}{5}\times 1467.48 = 47.724$$

于是统计量 F 的值为

$$\frac{S_A^2/(4-1)}{S_e^2/4(5-1)} = \frac{134.448/3}{47.724/16} = 15.04$$

查 F 分布表，$\alpha = 0.01$ 时，$F_{0.01}(3,16) = 5.29$，现在 $15.04 > 5.29$ 说明不同品种间的平均产量有显著差异。

2　设各小学五年级男学生的身高 Y_1，Y_2，Y_3 相互独立且服从相同方差的正态分布 $N(\mu_i, \sigma^2)$，$i = 1,2,3$。

要求检验假设 H_0：$\mu_1 = \mu_2 = \mu_3$。

将题给数据列入下表进行计算

	一	二	三	左边各数相加
	128.1	150.3	140.6	
	134.1	147.9	143.1	
	133.1	136.8	144.5	
	138.9	126.0	143.7	
	140.8	150.7	148.5	
	127.4	155.8	146.4	
$\sum_j y_{ij}$	802.4	867.5	866.8	2536.7
$(\sum_j y_{ij})^2$	643845.76	752556.25	751342.24	2147744.25
$\sum_j y_{ij}^2$	107456.64	126038.87	125261.12	358756.63

由上表可得

$$S_A^2 = \frac{1}{6}\sum_{i=1}^{3}\left(\sum_{j=1}^{6} y_{ij}\right)^2 - \frac{1}{3\times 6}\left(\sum_{i=1}^{3}\sum_{j=1}^{6} y_{ij}\right)^2$$

$$= \frac{1}{6}\times 2147744.25 - \frac{1}{18}(2536.7)^2 = 465.8812$$

$$S_E^2 = \sum_{i=1}^{3}\sum_{j=1}^{6} y_{ij}^2 - \frac{1}{6}\sum_{i=1}^{3}\left(\sum_{j=1}^{6} y_{ij}\right)^2$$

$$= 358756.63 - \frac{1}{6}\times 2147744.25 = 799.2550$$

于是，

$$F = \frac{465.8812/(3-1)}{799.2550/3(6-1)} = 4.3717$$

给定$\alpha = 0.05$，查F分布表得$F_{0.05}(2,15) = 3.68$．现在$4.3717 > 3.68$所以拒绝H_0，即认为三所小学五年级男学生的身高有显著差别．

3　将题给数据均减去18后列成下表．

机器A	操作工B	数据			合计
M_1	1	−3	−3	−1	−7
M_1	2	1	1	−2	0
M_1	3	−2	0	3	1
M_2	1	−1	−1	−1	−3
M_2	2	−3	−3	−3	−9
M_2	3	1	4	4	+9
M_3	1	−3	−1	−2	−6
M_3	2	0	−1	−2	−3
M_3	3	0	0	0	0
M_4	1	0	2	4	6
M_4	2	−3	−2	−1	−6
M_4	3	−1	−1	−1	−3
					$\Sigma = -21$

各列对应"1"的"合计"之和	−6	−10
各列对应"2"的"合计"之和	−3	−18
各列对应"3"的"合计"之和	−9	+7
各列对应"4"的"合计"之和	−3	

由上表可得

$$CT' = \frac{1}{36}(-21)^2 = \frac{49}{4} = 12.25$$

$$S_{T'}^2 = (-3)^2 + (-3)^2 + (-1)^2 + \cdots + (-1)^2 - \frac{49}{4}$$

$$= 144.75$$

$$S_A^2 = \frac{1}{3\times 3}[(-6)^2 + (-3)^2 + (-9)^2 + (-3)^2] - 12.25$$

$$= 2.75$$

$$S_B^2 = \frac{1}{4\times 3}[(-10)^2 + (-18)^2 + 7^2] - 12.25 = 27.167$$

$$S_{A\times B}^2 = \frac{1}{3}[(-7)^2 + 0^2 + 1^2 + \cdots + (-6)^2 + (-3)^2] - S_A^2$$

$$- S_B^2 - CT' = 73.50$$

$$S_e^2 = (-3)^2 + (-3)^2 + \cdots + (-1)^2 - \frac{1}{3}[(-7)^2 + 0^2 +$$

$$\cdots + (-3)^2] = 157 - 115.667 = 41.333$$

于是，

$$F_A = \frac{2.75/(4-1)}{41.333/4\times 3\times(3-1)} = \frac{0.92}{1.72} = 0.53$$

$$F_B = \frac{27.167/(3-1)}{41.333/4\times 3\times(3-1)} = \frac{13.58}{1.72} = 7.89$$

$$F_{A\times B} = \frac{73.50/(4-1)\times(3-1)}{41.333/4\times 3\times(3-1)} = \frac{12.25}{1.72} = 7.11$$

查 F 分布表知（取$\alpha = 0.05$）

$F_{0.05}(3,24) = 3.01$，$F_{0.05}(2,24) = 3.40$，$F_{0.05}(6,24) = 2.51$，由于$0.53 < 3.01$，$7.89 > 3.40$，$7.11 > 2.51$知机器之间的差别不显著，而操作工之间的差异及交互作用均是显著的。

4　将题给数据均减去43后列成下表。

苗别(A)	株行距(B)	数 据		合 计
1	1	2.4	5.0	7.4
1	2	4.0	2.0	6.0
1	3	0	2.2	2.2
1	4	4.6	6.2	10.8
2	1	1.1	−2.9	−1.8
2	2	−1.3	−5.5	−6.8
2	3	−6.1	−2.1	−8.2
2	4	−7.0	−2.8	−9.8
各列对应“1”的“合计”之和	26.4	5.6		Σ = −0.2
各列对应“2”的“合计”之和	−26.6	−0.8		
各列对应“3”的“合计”之和		−6.0		
各列对应“4”的“合计”之和		1		

由上表可得

$$CT' = \frac{1}{16}(-0.2)^2 = \frac{1}{400} = 0.0025$$

$$S_{T'}^2 = (2.4)^2 + (5.0)^2 + \cdots + (-7.0)^2 + (-2.8)^2 - CT'$$
$$= 255.2175$$

$$S_A^2 = \frac{1}{4\times 2}[26.4^2 + (-26.6)^2] - CT' = 175.5625$$

$$S_B^2 = \frac{1}{2\times 2}[5.6^2 + (-0.8)^2 + (-6.0)^2 + 1^2] - CT'$$
$$= 17.2475$$

$$S_{A\times B}^2 = \frac{1}{2}[7.4^2 + 6.0^2 + \cdots + (-9.8)^2] - S_A^2 - S_B^2 - CT'$$
$$= 19.6875$$

$$S_e^2 = (2.4)^2 + (5.0)^2 + \cdots + (-2.8)^2 - \frac{1}{2}[7.4^2 + 6.0^2$$

$$+\cdots+(-9.8)^2]=42.72$$

于是，

$$F_A=\frac{175.5625/1}{42.72/2\times4\times(2-1)}=\frac{175.5625}{42.72/8}=32.88$$

$$F_B=\frac{17.2475/(4-1)}{42.72/2\times4\times(2-1)}=1.08$$

$$F_{A\times B}=\frac{19.6875/(2-1)(4-1)}{42.72/2\times4\times(2-1)}=1.23$$

查 F 分布表知（取$\alpha=0.01$），

$$F_{0.01}(1,8)=11.26,\ F_{0.01}(3,8)=7.59$$

由于32.88＞11.26，1.08＜7.59，1.23＜7.59知不同育苗方式之间的差异是高度显著的，而株行距之间的差异与交互作用均不显著。

5　由题设数据将主要计算结果列出

$\sum x=30$	$\sum y=1052$	$n=5$
$\bar{x}=6$	$\bar{y}=210.4$	
$\sum x^2=220$	$\sum y^2=275990$	
$(\sum x)^2/n=180$	$(\sum y)^2/n=221340.8$	
$\sum xy=7790$	$(\sum x)(\sum y)/n=6312$	

$$\hat{\beta}=\frac{L_{XY}}{L_{XX}}=\frac{\sum xy-(\sum x)(\sum y)/n}{\sum x^2-(\sum x)^2/n}=36.95$$

$$\hat{\alpha}=\bar{y}-\hat{\beta}\,\bar{x}=-11.3$$

欲检验假设H_0：$\beta=0$．考虑统计量

$$\mathrm{F}=\frac{U}{Q/(n-2)}=\frac{\hat{\beta}\,\mathrm{L}_{XY}}{(\mathrm{L}_{YY}-U)/(n-2)}$$

现在，

$$\hat{\beta}=36.95$$

$$\mathrm{L}_{XY}=7790-6312=1478$$

$$U=36.95\times1478=54612.1$$

$$L_{yy} = 275990 - 221340.8 = 54649.2$$

$$Q = L_{yy} - U = 37.1$$

从而

$$F = \frac{54612.1}{37.1/(5-2)} = 4416.07$$

给定$\alpha = 0.01$，查F分布表得$F_{0.01}(1,3) = 34.12$，由$4416.07 > 34.12$，故否定H_0，即认为β显著地不为0。

6　由题设数据，将主要计算结果列出

$\sum x = 127.5$	$\sum y = 113.1$	$n = 24$
$\bar{x} = 5.31$	$\bar{y} = 4.71$	
$\sum x^2 = 829.61$	$\sum y^2 = 650.93$	
$(\sum x)^2/n = 677.34$	$(\sum y)^2/n = 532.98$	
$\sum xy = 731.60$	$(\sum x)(\sum y)/n = 600.84$	

$$\hat{\beta} = \frac{731.60 - 600.84}{829.61 - 677.34} = \frac{130.76}{152.27} = 0.859$$

$$\hat{\alpha} = 4.71 - 0.859 \times 5.31 = 0.15$$

因此所求之回归直线方程为

$$\hat{Y} = 0.15 + 0.859x$$

样本相关系数

$$r = \frac{L_{xy}}{\sqrt{L_{xx}L_{yy}}} = \frac{130.76}{\sqrt{152.27 \times 117.95}} = 0.976$$

查相关系数检验表，对$\alpha = 0.01$，$n - 2 = 24 - 2 = 22$，临界值为0.515，现在$0.976 > 0.515$，所以Y与x之间的线性关系高度显著。

7　由题设数据，令　$x' = x - 30$，$y' = y - 3000$，将主要计算结果列出

$\sum x'^2 = 84.37$	$\sum y'^2 = 4726421.05$	$n = 9$
$(\sum x')^2/n = 30.99$	$(\sum y')^2/n = 125009.387$	
$L_{xx} = 53.38$	$L_{yy} = 4601411.663$	

$$\sum x'y' = 12255.01$$
$$(\sum x')(\sum y')/n = -1968.188$$
$$L_{xy} = 14223.198$$

样本相关系数

$$r = \frac{L_{xy}}{\sqrt{L_{xx}L_{yy}}} = \frac{14223.198}{15672.376} = 0.9075$$

查相关系数检验表，对$\alpha = 0.05$，$n-2=7$，临界值是0.666，现在0.9075>0.666，因此Y与X的线性相关关系显著。

8　将数据在直角坐标系里点出，可看见菜青虫的个数随着停药后的天数增加急剧上升，因此，我们可以考虑选用指数函数类型的曲线

$$\hat{Y} = AB^x$$

来表示菜青虫的个数与喷药后天数之间的关系，其中A、B均为待定之常数。

令$Y' = \lg Y$，$\alpha = \lg A$，$\beta = \lg B$　于是有

$$\hat{Y}' = \alpha + \beta x$$

这样我们就可用一元线性回归的办法定出系数α，β。将此题的数据列出

x	0	1	2	3	4	5	6	7	8	9	10	11
y	23	28	33	40	48	57	68	81	98	116	142	167
$y' = \lg y$	1.362	1.447	1.519	1.602	1.681	1.756	1.833	1.909	1.991	2.065	2.152	2.223

主要的计算结果为

$\sum x = 66$	$\sum y' = 21.54$	$n = 12$
$\bar{x} = 5.5$	$\bar{y}' = 1.795$	$\sum xy' = 129.642$
$\sum x^2 = 506$	$\sum y'^2 = 39.537$	$(\sum x)(\sum y')/n = 118.47$
$(\sum x)^2/n = 363$	$(\sum y')^2/n = 38.664$	

$$\hat{\beta} = \frac{129.642 - 118.47}{506 - 363} = \frac{11.172}{143} = 0.07813$$

$$\hat{\alpha} = \overline{y}' - \hat{\beta}\overline{x} = 1.795 - 0.07813 \times 5.5 = 1.3653$$

因此

$$\hat{Y}' = 1.3653 + 0.07813x$$

再由　$A = 10^{\alpha}$，$B = 10^{\beta}$　定出$A = 23.19$，$B = 1.197$，于是，

$$\hat{Y} = AB^x = 23.19 \times 1.197^x$$

即为所求之菜青虫的个数（Y）与喷药后天数（x）之间的关系式。

9

（1）
$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 1 & 2 \end{pmatrix} \qquad \beta = \begin{pmatrix} a \\ b \end{pmatrix}$$

（2）

$$C'C = \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & -1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 5 \end{pmatrix}$$

（3）正规方程组为$(C'C)\hat{\beta} = C'Y$，即

$$\begin{pmatrix} 6 & 0 \\ 0 & 5 \end{pmatrix} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} Y_1 + 2Y_2 + Y_3 \\ -Y_2 + 2Y_3 \end{pmatrix}$$

（4）β的最小二乘估计$\hat{\beta}$为

$$\hat{\beta} = \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 5 \end{pmatrix}^{-1} \begin{pmatrix} Y_1 + 2Y_2 + Y_3 \\ -Y_2 + 2Y_3 \end{pmatrix}$$

$$= \begin{pmatrix} \frac{1}{6} & 0 \\ 0 & \frac{1}{5} \end{pmatrix} \begin{pmatrix} Y_1 + 2Y_2 + Y_3 \\ -Y_2 + 2Y_3 \end{pmatrix}$$

$$
=\begin{pmatrix} \frac{1}{6}(Y_1+2Y_2+Y_3) \\ \\ \frac{1}{5}(\quad -Y_2+2Y_3) \end{pmatrix}
$$

(5) $\hat{\beta}$ 的方差与协方差矩阵 $D(\hat{\beta})$ 为

$$
D(\hat{\beta})=\sigma^2(C'C)^{-1}
$$

$$
=\sigma^2\begin{pmatrix} \frac{1}{6} & 0 \\ 0 & \frac{1}{5} \end{pmatrix}
$$

从而

$$
D(\hat{a})=\frac{1}{6}\sigma^2, \qquad D(\hat{b})=\frac{1}{5}\sigma^2
$$

10 这一天平称重问题可写为下述矩阵形式：

$$
\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix}=\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \end{pmatrix}+\begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \end{pmatrix}
$$

其中 ε_1，ε_2，ε_3，ε_4 为测量误差，可假定它们是相互独立的随机变量，且$E(\varepsilon_i)=0$，$D(\varepsilon_i)=\sigma^2$，$i=1,2,3,4$.

这是一个线性模型．系数阵

$$
C=\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}
$$

由此 $\hat{\beta}$ 应满足正规方程 $(C'C)\hat{\beta}=C'Y$，即

$$
\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \hat{\beta}_4 \end{pmatrix}
$$

$$
=\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \end{pmatrix}
$$

$$
\begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} \begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \\ \hat{\beta}_4 \end{pmatrix} = \begin{pmatrix} Y_1 + Y_2 + Y_3 + Y_4 \\ Y_1 - Y_2 - Y_3 - Y_4 \\ Y_1 + Y_2 - Y_3 - Y_4 \\ Y_1 - Y_2 - Y_3 + Y_4 \end{pmatrix}
$$

因此，$\hat{\beta}_1 = \frac{1}{4}(Y_1 + Y_2 + Y_3 + Y_4)$

$$\hat{\beta}_2 = \frac{1}{4}(Y_1 - Y_2 + Y_3 - Y_4)$$

$$\hat{\beta}_3 = \frac{1}{4}(Y_1 + Y_2 - Y_3 - Y_4)$$

$$\hat{\beta}_4 = \frac{1}{4}(Y_1 - Y_2 - Y_3 + Y_4)$$

由题给的数据代入计算得

$$\hat{\beta}_1 = 9.7, \qquad \hat{\beta}_2 = 5, \qquad \hat{\beta}_3 = 4.4, \qquad \hat{\beta}_4 = 1.1$$

附　　表

表 I　　函数$\frac{\lambda^k}{k!}e^{-\lambda}$数值表

λ / k	0.1	0.2	0.3	0.4	0.5	0.6
0	0.90484	0.81873	0.74082	0.67032	0.60653	0.54881
1	0.09048	0.16375	0.22225	0.26813	0.30327	0.32929
2	0.00452	0.01637	0.03334	0.05362	0.07581	0.09878
3	0.00015	0.00109	0.00333	0.00715	0.01263	0.01976
4	—	0.00005	0.00025	0.00071	0.00158	0.00296
5	—	—	0.00001	0.00005	0.00016	0.00036
6	—	—	—	—	0.00001	0.00004
7	—	—	—	—	—	—
8	—	—	—	—	—	—
9	—	—	—	—	—	—
10	—	—	—	—	—	—
11	—	—	—	—	—	—
12	—	—	—	—	—	—
13	—	—	—	—	—	—
14	—	—	—	—	—	—
15	—	—	—	—	—	—
16	—	—	—	—	—	—
17	—	—	—	—	—	—
18	—	—	—	—	—	—
19	—	—	—	—	—	—
20	—	—	—	—	—	—
21	—	—	—	—	—	—
22	—	—	—	—	—	—
23	—	—	—	—	—	—
24	—	—	—	—	—	—
25	—	—	—	—	—	—
26	—	—	—	—	—	—
27	—	—	—	—	—	—
28	—	—	—	—	—	—
29	—	—	—	—	—	—
30	—	—	—	—	—	—
31	—	—	—	—	—	—
32	—	—	—	—	—	—
33	—	—	—	—	—	—
34	—	—	—	—	—	—

续表 I

k \ λ	0.7	0.8	0.9	1	2	3
0	0.49659	0.44933	0.40657	0.36788	0.13534	0.04978
1	0.34761	0.35946	0.36591	0.36788	0.27067	0.14936
2	0.12166	0.14379	0.16466	0.18394	0.27067	0.22404
3	0.02838	0.03834	0.04939	0.06131	0.18045	0.22044
4	0.00496	0.00766	0.01111	0.01532	0.09022	0.16803
5	0.00069	0.00123	0.00200	0.00306	0.03609	0.10082
6	0.00008	0.00016	0.00030	0.00051	0.01203	0.05040
7	—	0.00001	0.00003	0.00007	0.00343	0.02160
8	—	—	—	—	0.00085	0.00810
9	—	—	—	—	0.00019	0.00270
10	—	—	—	—	0.00003	0.00081
11	—	—	—	—	—	0.00022
12	—	—	—	—	—	0.00005
13	—	—	—	—	—	0.00001
14	—	—	—	—	—	—
15	—	—	—	—	—	—
16	—	—	—	—	—	—
17	—	—	—	—	—	—
18	—	—	—	—	—	—
19	—	—	—	—	—	—
20	—	—	—	—	—	—
21	—	—	—	—	—	—
22	—	—	—	—	—	—
23	—	—	—	—	—	—
24	—	—	—	—	—	—
25	—	—	—	—	—	—
26	—	—	—	—	—	—
27	—	—	—	—	—	—
28	—	—	—	—	—	—
29	—	—	—	—	—	—
30	—	—	—	—	—	—
31	—	—	—	—	—	—
32	—	—	—	—	—	—
33	—	—	—	—	—	—
34	—	—	—	—	—	—

续表 I

k \ λ	4	5	6	7	8	9
0	0.01831	0.00673	0.00247	0.00091	0.00033	0.00012
1	0.07326	0.03369	0.01487	0.00638	0.00268	0.00111
2	0.14653	0.08422	0.04461	0.02234	0.01073	0.00499
3	0.19537	0.14037	0.08923	0.05212	0.02862	0.01499
4	0.19537	0.17547	0.13385	0.09122	0.05725	0.03373
5	0.15629	0.17547	0.16062	0.12772	0.09160	0.06072
6	0.10420	0.14622	0.16062	0.14900	0.12214	0.09109
7	0.05954	0.10444	0.13768	0.14900	0.13950	0.11712
8	0.02977	0.06527	0.10326	0.13038	0.13959	0.13176
9	0.01323	0.03626	0.06883	0.10140	0.12408	0.13176
10	0.00529	0.01813	0.04130	0.07098	0.09926	0.11858
11	0.00192	0.00824	0.02252	0.04517	0.07219	0.09702
12	0.00064	0.00343	0.01126	0.02635	0.04812	0.07276
13	0.00019	0.00132	0.00519	0.01418	0.02961	0.05037
14	0.00005	0.00047	0.00222	0.00709	0.01692	0.03238
15	0.00001	0.00015	0.00089	0.00331	0.00902	0.01943
16	—	0.00004	0.00033	0.00144	0.00451	0.01093
17	—	0.00001	0.00011	0.00059	0.00212	0.00578
18	—	—	0.00003	0.00023	0.00094	0.00289
19	—	—	0.00001	0.00008	0.00039	0.00137
20	—	—	—	0.00003	0.00015	0.00061
21	—	—	—	—	0.00006	0.00026
22	—	—	—	—	0.00002	0.00010
23	—	—	—	—	—	0.00004
24	—	—	—	—	—	0.00001
25	—	—	—	—	—	—
26	—	—	—	—	—	—
27	—	—	—	—	—	—
28	—	—	—	—	—	—
29	—	—	—	—	—	—
30	—	—	—	—	—	—
31	—	—	—	—	—	—
32	—	—	—	—	—	—
33	—	—	—	—	—	—
34	—	—	—	—	—	—

续表 I

k \ λ	10	11	12	13	14	15
0	0.00004	0.00001	—	—	—	—
1	0.00045	0.00018	0.00007	0.00002	0.00001	—
2	0.00227	0.00101	0.00044	0.00019	0.00008	0.00003
3	0.00756	0.00370	0.00177	0.00082	0.00038	0.00017
4	0.01891	0.01018	0.00530	0.00269	0.00133	0.00064
5	0.03783	0.02241	0.01274	0.00699	0.00373	0.00193
6	0.06305	0.04109	0.02548	0.01515	0.00869	0.00483
7	0.09007	0.06457	0.04368	1.02814	0.01739	0.01037
8	0.11260	0.08879	0.06552	0.04573	0.03043	0.01944
9	0.12511	0.10853	0.08736	0.06605	0.04734	0.03240
10	0.12511	0.11938	0.10484	0.08587	0.06628	0.04861
11	0.11374	0.11938	0.11437	0.10148	0.08435	0.06628
12	0.09478	0.10943	0.11437	0.10994	0.09841	0.08285
13	0.07290	0.09259	0.10557	0.10994	0.10599	0.09560
14	0.05207	0.07275	0.09048	0.10209	0.10599	0.10244
15	0.03471	0.05335	0.07239	0.08847	0.09892	0.10244
16	0.02169	0.03668	0.05429	0.07188	0.08655	0.09603
17	0.01276	0.02373	0.03832	0.05497	0.07128	0.08473
18	0.00709	0.01450	0.02555	0.03970	0.05544	0.07061
19	0.00373	0.00839	0.01613	0.02716	0.04085	0.05574
20	0.00186	0.00461	0.00968	0.01765	0.02859	0.04181
21	0.00088	0.00241	0.00553	0.01093	0.01906	0.02986
22	0.00040	0.00121	0.00301	0.00645	0.01213	0.02036
23	0.00017	0.00057	0.00157	0.00365	0.00738	0.01328
24	0.00007	0.00026	0.00078	0.00197	0.00430	0.00830
25	0.00002	0.00011	0.00037	0.00102	0.00241	0.00498
26	0.00001	0.00004	0.00017	0.00051	0.00129	0.00287
27	—	0.00002	0.00007	0.00024	0.00067	0.00159
28	—	—	0.00003	0.00011	0.00033	0.00085
29	—	—	0.00001	0.00005	0.00016	0.00044
30	—	—	—	0.00002	0.00007	0.00022
31	—	—	—	—	0.00003	0.00010
32	—	—	—	—	0.00001	0.00005
33	—	—	—	—	—	0.00002
34	—	—	—	—	—	0.00001

表 I　函数 $\sum_{k=n}^{\infty} \frac{\lambda^k}{k!} e^{-\lambda}$ 数值表

n \ λ	0.1	0.2	0.3	0.4	0.5	0.6
1	0.095163	0.181269	0.259182	0.329680	0.393469	0.451188
2	0.004679	0.017523	0.036936	0.061552	0.090204	0.121901
3	0.000155	0.001148	0.003600	0.007926	0.014388	0.023115
4	0.000004	0.000057	0.000266	0.000776	0.001752	0.003358
5	—	0.000002	0.000016	0.000061	0.000172	0.000394
6	—	—	0.000001	0.000004	0.000014	0.000038
7	—	—	—	0.000000	0.000001	0.000003
8	—	—	—	—	—	—
9	—	—	—	—	—	—
10	—	—	—	—	—	—
11	—	—	—	—	—	—
12	—	—	—	—	—	—
13	—	—	—	—	—	—
14	—	—	—	—	—	—
15	—	—	—	—	—	—
16	—	—	—	—	—	—
17	—	—	—	—	—	—
18	—	—	—	—	—	—
19	—	—	—	—	—	—
20	—	—	—	—	—	—
21	—	—	—	—	—	—
22	—	—	—	—	—	—
23	—	—	—	—	—	—
24	—	—	—	—	—	—
25	—	—	—	—	—	—
26	—	—	—	—	—	—

续表 Ⅱ

n \ λ	0.7	0.8	0.9	1	2	3
1	0.503415	0.550671	0.593430	0.63212	0.86466	0.95021
2	0.155805	0.191208	0.227517	0.26424	0.59399	0.80085
3	0.034142	0.047423	0.062856	0.08030	0.32332	0.57681
4	0.005754	0.009080	0.011458	0.01899	0.14288	0.35277
5	0.000786	0.001411	0.002343	0.00366	0.05265	0.18474
6	0.000091	0.000184	0.000342	0.00059	0.01656	0.08392
7	0.000010	0.000020	0.000042	0.00008	0.00453	0.03351
8	0.000002	0.000001	0.000005	0.00001	0.00110	0.01191
9	—	—	—	0.00000	0.00024	0.00380
10	—	—	—	—	0.00005	0.00110
11	—	—	—	—	0.00001	0.00029
12	—	—	—	—	—	0.00007
13	—	—	—	—	—	0.00002
14	—	—	—	—	—	—
15	—	—	—	—	—	—
16	—	—	—	—	—	—
17	—	—	—	—	—	—
18	—	—	—	—	—	—
19	—	—	—	—	—	—
20	—	—	—	—	—	—
21	—	—	—	—	—	—
22	—	—	—	—	—	—
23	—	—	—	—	—	—
24	—	—	—	—	—	—
25	—	—	—	—	—	—
26	—	—	—	—	—	—

续表 I

n \ λ	4	5	6	7	8	9	10
1	0.981680	0.99326	0.99752	0.99909	0.99966	0.99988	0.99996
2	0.90842	0.95957	0.98265	0.99271	0.99698	0.99877	0.99950
3	0.76190	0.87535	0.93803	0.97036	0.98625	0.99377	0.99724
4	0.56653	0.73497	0.84880	0.91823	0.95762	0.97877	0.98966
5	0.37116	0.55951	0.71494	0.82701	0.90037	0.94504	0.97075
6	0.21487	0.38404	0.55432	0.69929	0.80876	0.88431	0.93291
7	0.11067	0.23782	0.39370	0.55029	0.68663	0.79322	0.86986
8	0.05113	0.13337	0.25602	0.40129	0.54704	0.67610	0.77978
9	0.02137	0.06809	0.15276	0.27091	0.40745	0.54435	0.66719
10	0.00813	0.03183	0.08392	0.16950	0.28338	0.41259	0.54207
11	0.00284	0.01369	0.04262	0.09852	0.18412	0.29401	0.41696
12	0.00092	0.00545	0.02009	0.05335	0.11192	0.19699	0.30322
13	0.00027	0.00202	0.00883	0.02700	0.06380	0.12423	0.20844
14	0.00008	0.00070	0.00363	0.01281	0.03418	0.07385	0.13554
15	0.000002	0.00023	0.00140	0.00572	0.01726	0.04147	0.08346
16	0.000001	0.00007	0.00051	0.00241	0.00823	0.02204	0.04874
17	—	0.00002	0.00017	0.00096	0.00372	0.01111	0.02704
18	—	0.00001	0.00006	0.00036	0.00160	0.00533	0.01428
19	—	—	0.00002	0.00013	0.00065	0.00243	0.00719
20	—	—	0.00001	0.00005	0.00025	0.00106	0.00345
21	—	—	—	—	0.00010	0.00044	0.00159
22	—	—	—	—	0.00003	0.00018	0.00070
23	—	—	—	—	0.00001	0.00006	0.00030
24	—	—	—	—	—	0.00003	0.00012
25	—	—	—	—	—	—	0.00004
26	—	—	—	—	—	—	0.00001

表Ⅱ　　函数 $\varphi(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$ 数值表

x	0	1	2	3	4	5	6	7	8	9
0.0	0.3989	3989	3989	3988	3986	3984	3982	3980	3977	3973
0.1	3970	3965	3961	3956	3951	3945	3939	3932	3925	3918
0.2	3910	3902	3894	3885	3876	3867	3857	3847	3836	3825
0.3	3814	3802	3790	3778	3765	3752	3739	3726	3712	3697
0.4	3683	3668	3653	3637	3621	3605	3589	3572	3555	3538
0.5	3521	3503	3485	3467	3448	3429	3411	3391	3372	3352
0.6	3332	3312	3292	3271	3251	3230	3209	3187	3166	3144
0.7	3123	3101	3079	3056	3034	3011	2989	2966	2943	2920
0.8	2897	2874	2850	2827	2803	2780	2756	2732	2709	2685
0.9	2661	2637	2613	2589	2565	2541	2516	2492	2468	2444
1.0	0.2420	2396	2371	2347	2323	2299	2275	2251	2227	2203
1.1	2179	2155	2131	2107	2083	2059	2036	2012	1989	1965
1.2	1942	1919	1895	1872	1849	1826	1804	1781	1759	1736
1.3	1714	1691	1669	1647	1626	1604	1582	1561	1540	1518
1.4	1497	1476	1456	1435	1415	1394	1374	1354	1334	1315
1.5	1295	1276	1257	1238	1219	1200	1182	1163	1145	1127
1.6	1109	1092	1074	1057	1040	1023	1006	0989	0973	0957
1.7	0940	0925	0909	0893	0878	0863	0848	0833	0818	0804
1.8	0790	0775	0761	0748	0734	0721	0707	0694	0681	0669
1.9	0656	0644	0632	0620	0608	0596	0584	0573	0562	0551
2.0	0.0540	0529	0519	0508	0498	0488	0478	0468	0459	0449
2.1	0440	0431	0422	0413	0404	0396	0387	0379	0371	0363
2.2	0355	0347	0339	0332	0325	0317	0310	0303	0297	0290
2.3	0283	0277	0270	0264	0258	0252	0246	0241	0235	0229
2.4	0224	0219	0213	0208	0203	0198	0194	0189	0184	0180
2.5	0175	0171	0167	0163	0158	0154	0151	0147	0143	0139
2.6	0136	0132	0129	0126	0122	0119	0116	0113	0110	0107
2.7	0104	0101	0099	0096	0093	0091	0088	0086	0084	0081
2.8	0079	0077	0075	0073	0071	0069	0067	0065	0063	0061
2.9	0060	0058	0056	0055	0053	0051	0050	0048	0047	0046
3.0	0.0044	0043	0042	0040	0039	0038	0037	0036	0035	0034
3.1	0033	0032	0031	0030	0029	0028	0027	0026	0025	0025
3.2	0024	0023	0022	0022	0021	0020	0020	0019	0018	0018
3.3	0017	0017	0016	0016	0015	0015	0014	0014	0013	0013
3.4	0012	0012	0012	0011	0011	0010	0010	0010	0009	0009
3.5	0009	0008	0008	0008	0008	0007	0007	0007	0007	0006
3.6	0006	0006	0006	0005	0005	0005	0005	0005	0005	0004
3.7	0004	0004	0004	0004	0004	0004	0003	0003	0003	0003
3.8	0003	0003	0003	0003	0003	0002	0002	0002	0002	0002
3.9	0002	0002	0002	0002	0002	0002	0002	0002	0001	0001

表Ⅳ　　函数 $= \Phi(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{x} e^{-\frac{z^2}{2}} dz$ 数值表

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.00	0.5000	0.35	0.6368	0.70	0.7580	1.05	0.8531
0.01	0.5040	0.36	0.6406	0.71	0.7611	1.06	0.8554
0.02	0.5080	0.37	0.6443	0.72	0.7642	1.07	0.8577
0.03	0.5120	0.38	0.6480	0.73	0.7673	1.08	0.8599
0.04	0.5160	0.39	0.6517	0.74	0.7703	1.09	0.8621
0.05	0.5199	0.40	0.6554	0.75	0.7734	1.10	0.8643
0.06	0.5239	0.41	0.6591	0.76	0.7764	1.11	0.8665
0.07	0.5279	0.42	0.6628	0.77	0.7794	1.12	0.8686
0.08	0.5319	0.43	0.6664	0.78	0.7823	1.13	0.8708
0.09	0.5359	0.44	0.6700	0.79	0.7852	1.14	0.8729
0.10	0.5398	0.45	0.6736	0.80	0.7881	1.15	0.8749
0.11	0.5438	0.46	0.6772	0.81	0.7910	1.16	0.8770
0.12	0.5478	0.47	0.6808	0.82	0.7939	1.17	0.8790
0.13	0.5517	0.48	0.6844	0.83	0.7967	1.18	0.8810
0.14	0.5557	0.49	0.6879	0.84	0.7995	1.19	0.8830
0.15	0.5596	0.50	0.6915	0.85	0.8023	1.20	0.8849
0.16	0.5636	0.51	0.6950	0.86	0.8051	1.21	0.8869
0.17	0.5675	0.52	0.6985	0.87	0.8078	1.22	0.8888
0.18	0.5714	0.53	0.7019	0.88	0.8106	1.23	0.8907
0.19	0.5753	0.54	0.7054	0.89	0.8133	1.24	0.8925
0.20	0.5793	0.55	0.7088	0.90	0.8159	1.25	0.8944
0.21	0.5832	0.56	0.7123	0.91	0.8186	1.26	0.8962
0.22	0.5871	0.57	0.7157	0.92	0.8212	1.27	0.8980
0.23	0.5910	0.58	0.7190	0.93	0.8238	1.28	0.8997
0.24	0.5948	0.59	0.7224	0.94	0.8264	1.29	0.9015
0.25	0.5987	0.60	0.7257	0.95	0.8289	1.30	0.9032
0.26	0.6026	0.61	0.7291	0.96	0.8315	1.31	0.9049
0.27	0.6064	0.62	0.7324	0.97	0.8340	1.32	0.9066
0.28	0.6103	0.63	0.7357	0.98	0.8365	1.33	0.9082
0.29	0.6141	0.64	0.7389	0.99	0.8389	1.34	0.9099
0.30	0.6179	0.65	0.7422	1.00	0.8413	1.35	0.9115
0.31	0.6217	0.66	0.7454	1.01	0.8438	1.36	0.9131
0.32	0.6255	0.67	0.7486	1.02	0.8461	1.37	0.9147
0.33	0.6293	0.68	0.7517	1.03	0.8485	1.38	0.9162
0.34	0.6331	0.69	0.7549	1.04	0.8508	1.39	0.9177

续表Ⅳ

x	Φ(x)	x	Φ(x)	x	Φ(x)	x	Φ(x)
1.40	0.9192	1.75	0.9599	2.20	0.9861	2.90	0.9981
1.41	0.9207	1.76	0.9608	2.22	0.9868	2.92	0.9982
1.42	0.9222	1.77	0.9616	2.24	0.9875	2.94	0.9984
1.43	0.9236	1.78	0.9625	2.26	0.9881	2.96	0.9985
1.44	0.9251	1.79	0.9633	2.28	0.9887	2.98	0.9986
1.45	0.9265	1.80	0.9641	2.30	0.9893	3.00	0.99865
1.46	0.9279	1.81	0.9649	2.32	0.9998	3.20	0.99931
1.47	0.9292	1.82	0.9656	2.34	0.9904	3.40	0.99966
1.48	0.9306	1.83	0.9664	2.36	0.9909	3.60	0.999841
1.49	0.9319	1.84	0.9671	2.38	0.9913	3.80	0.999928
1.50	0.9332	1.85	0.9678	2.40	0.9918	4.00	0.999968
1.51	0.9345	1.86	0.9686	2.42	0.9922	4.50	0.999997
1.52	0.9357	1.87	0.9693	2.44	0.9927	5.00	0.99999997
1.53	0.9370	1.88	0.9699	2.46	0.9931		
1.54	0.9382	1.89	0.9706	2.48	0.9934		
1.55	0.9394	1.90	0.9713	2.50	0.9938		
1.56	0.9406	1.91	0.9719	2.52	0.9941		
1.57	0.9418	1.92	0.9726	2.54	0.9945		
1.58	0.9429	1.93	0.9732	2.56	0.9948		
1.59	0.9441	1.94	0.9738	2.58	0.9951		
1.60	0.9452	1.95	0.9744	2.60	0.9953		
1.61	0.9463	1.96	0.9750	2.62	0.9956		
1.62	0.9474	1.97	0.9756	2.64	0.9959		
1.63	0.9484	1.98	0.9761	2.66	0.9961		
1.64	0.9495	1.99	0.9767	2.68	0.9963		
1.65	0.9505	2.00	0.9772	2.70	0.9965		
1.66	0.9515	2.02	0.9783	2.72	0.9967		
1.67	0.9525	2.04	0.9793	2.74	0.9969		
1.68	0.9535	2.06	0.9803	2.76	0.9971		
1.69	0.9545	2.08	0.9812	2.78	0.9973		
1.70	0.9554	2.10	0.9821	2.80	0.9974		
1.71	0.9564	2.12	0.9830	2.82	0.9976		
1.72	0.9573	2.14	0.9838	2.84	0.9977		
1.73	0.9582	2.16	0.9846	2.86	0.9979		
1.74	0.9591	2.18	0.9854	2.88	0.9980		

χ^2分布的上侧分位数(χ^2_α)表

表V　$P(\chi^2_f > \chi^2_\alpha) = \alpha$

f \ α	0.99	0.98	0.95	0.90	0.80	0.70	0.50
1	0.0^3157	0.0^3628	0.0^2393	0.0158	0.0642	0.148	0.455
2	0.0201	0.0404	0.103	0.211	0.446	0.713	1.386
3	0.115	0.185	0.352	0.584	1.005	1.424	2.366
4	0.297	0.429	0.711	1.064	1.649	2.195	3.357
5	0.554	0.752	1.145	1.610	2.343	3.000	4.351
6	0.872	1.134	1.635	2.204	3.070	3.828	5.348
7	1.239	1.564	2.167	2.833	3.822	4.671	6.346
8	1.646	2.032	2.733	3.490	4.594	5.527	7.344
9	2.088	2.532	3.325	4.168	5.380	6.393	8.343
10	2.558	3.059	3.940	4.865	6.179	7.267	9.342
11	3.053	3.609	4.575	5.578	6.989	8.148	10.341
12	3.571	4.178	5.226	6.304	7.807	9.034	11.340
13	4.107	4.765	5.892	7.042	8.634	9.926	12.340
14	4.660	5.368	6.571	7.790	9.467	10.821	13.339
15	5.229	5.985	7.261	8.547	10.307	11.721	14.339
16	5.812	6.614	7.962	9.312	11.152	12.624	15.338
17	6.408	7.255	8.672	10.085	12.002	13.531	16.338
18	7.015	7.906	9.390	10.865	12.857	14.440	17.338
19	7.633	8.567	10.117	11.651	13.716	15.352	18.338
20	8.260	9.237	10.851	12.443	14.578	16.266	19.337
21	8.897	9.915	11.591	13.240	15.445	17.182	20.337
22	9.542	10.600	12.338	14.041	16.314	18.101	21.337
23	10.196	11.293	13.091	14.848	17.187	19.021	22.337
24	10.856	11.992	13.848	15.659	18.062	19.943	23.337
25	11.524	12.697	14.611	16.473	18.940	20.867	24.337
26	12.198	13.409	15.379	17.292	19.820	21.792	25.336
27	12.878	14.125	16.151	18.114	20.703	22.719	26.336
28	13.565	14.847	16.928	18.939	21.588	23.647	27.336
29	14.256	15.574	17.708	19.768	22.475	24.577	28.336
30	14.953	16.306	18.493	20.599	23.364	25.508	29.336

续表 V

f \ α	0.30	0.20	0.10	0.05	0.02	0.01	0.001
1	1.074	1.642	2.706	3.841	5.412	6.635	10.828
2	2.408	3.219	4.605	5.991	7.824	9.210	13.816
3	3.665	4.642	6.251	7.815	9.837	11.345	16.266
4	4.878	5.989	7.779	9.488	11.668	13.277	18.467
5	6.064	7.289	9.236	11.070	13.388	15.086	20.515
6	7.231	8.558	10.645	12.592	15.033	16.812	22.458
7	8.383	9.803	12.017	14.067	16.622	18.475	24.322
8	9.524	11.030	13.362	15.507	18.168	20.090	26.125
9	10.656	12.242	14.684	16.919	19.679	21.666	27.877
10	11.781	13.442	15.987	18.307	21.161	23.209	29.588
11	12.899	14.631	17.275	19.675	22.618	24.725	31.264
12	14.011	15.812	18.549	21.026	24.054	26.217	32.909
13	15.119	16.985	19.812	22.362	25.472	27.688	34.528
14	16.222	18.151	21.064	23.685	26.873	29.141	36.123
15	17.322	19.311	22.307	24.996	28.259	30.578	37.697
16	18.418	20.465	23.542	26.296	29.633	32.000	39.252
17	19.511	21.615	24.769	27.587	30.995	33.409	40.790
18	20.601	22.760	25.989	28.869	32.346	34.805	42.312
19	21.689	23.900	27.204	30.144	33.687	36.191	43.820
20	22.775	25.038	28.412	31.410	35.020	37.566	45.315
21	23.858	26.171	29.615	32.671	36.343	38.932	46.797
22	24.939	27.301	30.813	33.924	37.659	40.289	48.268
23	26.018	28.429	32.007	35.172	38.968	41.638	49.728
24	27.096	29.553	33.196	36.415	40.270	42.980	51.179
25	28.172	30.675	34.382	37.652	41.566	44.314	52.618
26	29.246	31.795	35.563	38.885	42.856	45.642	54.052
27	30.319	32.912	36.741	40.113	44.140	46.963	55.476
28	31.391	34.027	37.916	41.337	45.419	48.278	56.893
29	32.461	35.139	39.087	42.557	46.693	49.588	58.301
30	33.530	36.250	40.256	43.773	47.962	50.892	59.703

表 Ⅵ

$$P\left(|t|\geqslant t_\alpha\right)=\frac{2\Gamma\left(\frac{n+1}{2}\right)}{\sqrt{n\pi}\,\Gamma\left(\frac{n}{2}\right)}\int_{t_\alpha}^{\infty}\frac{1}{\left(1+\frac{t^2}{n}\right)^{\frac{n+1}{2}}}dt=\alpha$$

的数值表

n	α							
	0.80	0.60	0.40	0.20	0.10	0.05	0.02	0.01
1	0.325	0.727	1.376	3.078	6.314	12.706	31.821	63.657
2	0.289	0.617	1.061	1.886	2.920	4.303	6.965	9.925
3	0.277	0.584	0.978	1.638	2.353	3.182	4.541	5.841
4	0.271	0.569	0.941	1.533	2.132	2.776	3.747	4.604
5	0.267	0.559	0.920	1.476	2.015	2.571	3.365	4.032
6	0.265	0.553	0.906	1.440	1.943	2.447	3.143	3.707
7	0.263	0.549	0.896	1.415	1.895	2.365	2.998	3.499
8	0.262	0.546	0.889	1.397	1.860	2.306	2.896	3.355
9	0.261	0.543	0.883	1.383	1.833	2.262	2.821	3.250
10	0.260	0.542	0.879	1.372	1.812	2.228	2.764	3.169
11	0.260	0.540	0.876	1.363	1.796	2.201	2.718	3.106
12	0.259	0.539	0.873	1.356	1.782	2.179	2.681	3.055
13	0.259	0.538	0.870	1.350	1.771	2.160	2.650	3.012
14	0.258	0.537	0.868	1.345	1.761	2.145	2.624	2.977
15	0.258	0.536	0.866	1.341	1.753	2.131	2.602	2.947
16	0.258	0.535	0.865	1.337	1.746	2.120	2.583	2.921
17	0.257	0.534	0.863	1.333	1.740	2.110	2.567	2.898
18	0.257	0.534	0.862	1.330	1.734	2.101	2.552	2.878
19	0.257	0.533	0.861	1.328	1.729	2.093	2.539	2.861
20	0.257	0.533	0.860	1.325	1.725	2.086	2.528	2.845
21	0.257	0.532	0.859	1.323	1.721	2.080	2.518	2.831
22	0.256	0.532	0.858	1.321	1.717	2.074	2.508	2.819
23	0.256	0.532	0.858	1.319	1.714	2.069	2.500	2.807
24	0.256	0.531	0.857	1.318	1.711	2.064	2.492	2.797
25	0.256	0.531	0.856	1.316	1.708	2.060	2.485	2.787
26	0.256	0.531	0.856	1.315	1.706	2.056	2.479	2.779
27	0.256	0.531	0.855	1.314	1.703	2.052	2.473	2.771
28	0.256	0.530	0.855	1.313	1.701	2.048	2.467	2.763
29	0.256	0.530	0.854	1.311	1.699	2.045	2.462	2.756
30	0.256	0.530	0.854	1.310	1.697	2.042	2.457	2.750
40	0.255	0.529	0.851	1.303	1.684	2.021	2.423	2.704
60	0.254	0.527	0.818	1.296	1.671	2.000	2.390	2.660
120	0.254	0.526	0.815	1.289	1.658	1.980	2.358	2.617
∞	0.253	0.524	0.842	1.282	1.645	1.960	2.326	2.576

表Ⅶ　F分布表($\alpha=0.05$)

N_1 \ N_2	1	2	3	4	5	6	7	8
1	161.4	199.5	215.7	224.6	230.2	234.0	236.8	238.9
2	18.51	19.00	19.16	19.25	19.30	19.33	91.35	91.37
3	10.13	9.55	9.28	9.12	9.01	8.94	8.89	8.85
4	7.71	6.94	6.59	6.39	6.26	6.16	6.09	6.04
5	6.61	5.79	5.41	5.19	5.05	4.95	4.88	4.82
6	5.99	5.14	4.76	4.54	4.39	4.28	4.21	4.15
7	5.59	4.74	4.35	4.12	3.97	3.87	3.79	3.73
8	5.32	4.46	4.07	3.84	3.69	3.58	3.50	3.44
9	5.12	4.26	3.86	3.63	3.48	3.37	3.29	3.23
10	4.96	4.10	3.71	3.48	3.33	3.22	3.14	3.07
11	4.84	3.98	3.59	3.36	3.20	3.09	3.01	2.95
12	4.75	3.89	3.49	3.26	3.11	3.00	2.91	2.85
13	1.67	3.81	3.41	3.18	3.03	2.92	2.83	2.77
14	4.60	3.74	3.34	3.11	2.96	2.85	2.76	2.70
15	4.54	3.68	3.29	3.06	2.90	2.79	2.71	2.64
16	4.49	3.63	3.24	3.01	2.85	2.74	2.66	2.59
17	4.45	3.59	3.20	2.96	2.81	2.70	2.61	2.55
18	4.41	3.55	3.16	2.93	2.77	2.66	2.58	2.51
19	4.38	3.52	3.13	2.90	2.74	2.63	2.54	2.48
20	4.35	3.49	3.10	2.87	2.71	2.60	2.51	2.45
21	4.32	3.47	3.07	2.84	2.68	2.57	2.49	2.42
22	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40
23	4.28	3.42	3.03	2.80	2.64	2.53	2.44	2.37
24	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36
25	4.24	3.39	2.99	2.76	2.60	2.49	2.40	2.34
30	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27
40	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18
60	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10
120	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02
∞	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94

续表Ⅶ　F分布表($\alpha=0.05$)

N_2 \ N_1	9	10	12	15	20	60	∞
1	240.5	241.9	243.9	245.9	248.0	252.2	254.3
2	19.38	19.40	19.41	19.43	19.45	19.48	91.50
3	8.81	8.79	8.74	8.70	8.66	8.57	8.53
4	6.00	5.96	5.91	5.86	5.80	5.69	5.63
5	4.77	4.74	4.68	4.62	4.56	4.43	4.36
6	4.10	4.06	4.00	3.94	3.87	3.74	3.67
7	3.68	3.64	3.57	3.51	3.44	3.30	3.23
8	3.39	3.35	3.28	3.22	3.15	3.01	2.93
9	3.18	3.14	3.07	3.01	2.94	2.79	2.71
10	3.02	2.98	2.91	2.85	2.77	2.62	2.54
11	2.90	2.85	2.79	2.72	2.65	2.49	2.40
12	2.80	2.75	2.69	2.62	2.54	2.38	2.30
13	2.71	2.67	2.60	2.53	2.46	2.30	2.21
14	2.65	2.60	2.53	2.46	2.39	2.22	2.13
15	2.59	2.54	2.48	2.40	2.33	2.16	2.07
16	2.54	2.49	2.42	2.35	2.28	2.11	2.01
17	2.49	2.45	2.38	2.31	2.23	2.06	1.96
18	2.46	2.41	2.34	2.27	2.19	2.02	1.92
19	2.42	2.38	2.31	2.23	2.16	1.98	1.88
20	2.39	2.35	2.28	2.20	2.12	1.95	1.84
21	2.37	2.32	2.25	2.18	2.10	1.92	1.81
22	2.34	2.30	2.23	2.15	2.07	1.89	1.78
23	2.32	2.27	2.20	2.13	2.05	1.86	1.76
24	2.30	2.25	2.18	2.11	2.03	1.84	1.73
25	2.28	2.24	2.16	2.09	2.01	1.82	1.71
30	2.21	2.16	2.09	2.01	1.93	1.74	1.62
40	2.12	2.08	2.00	1.92	1.84	1.64	1.51
60	2.04	1.99	1.92	1.84	1.75	1.53	1.39
120	1.96	1.91	1.83	1.75	1.66	1.43	1.25
∞	1.88	1.83	1.75	1.67	1.57	1.32	1.00

续表Ⅶ　F分布表($\alpha=0.01$)

N_2 \ N_1	1	2	3	4	5	6	7	8
1	4052	4999.5	5403	5625	5764	5859	5928	5982
2	98.50	99.00	99.17	99.25	99.30	99.33	99.36	99.37
3	34.12	34.82	29.46	28.71	28.24	27.91	27.67	27.49
4	21.20	18.00	16.69	15.98	15.52	15.21	14.98	14.80
5	16.26	13.27	12.06	11.39	10.97	10.67	10.46	10.29
6	13.75	10.92	9.78	9.15	8.75	8.47	8.26	8.10
7	12.35	9.55	8.45	7.85	7.46	7.19	6.99	6.84
8	11.26	8.65	7.59	7.01	6.63	6.37	6.18	6.03
9	10.56	8.02	6.99	6.42	6.06	5.80	5.61	5.47
10	10.04	7.56	6.55	5.99	5.64	5.39	5.20	5.06
11	9.65	7.21	6.22	5.67	5.32	5.07	4.89	4.74
12	9.33	6.93	5.95	5.41	5.06	4.82	4.64	4.50
13	9.07	6.70	5.74	5.21	4.86	4.62	4.44	4.30
14	8.86	6.51	5.56	5.04	4.69	4.46	4.28	4.14
15	8.68	6.36	5.42	4.89	4.56	4.32	4.14	4.00
16	8.53	6.23	5.29	4.77	4.44	4.20	4.03	3.89
17	8.40	6.11	5.18	4.67	4.34	4.10	3.93	3.79
18	8.29	6.01	5.09	4.58	4.25	4.01	3.84	3.71
19	8.18	5.93	5.01	4.50	4.17	3.94	3.77	3.63
20	8.10	5.85	4.94	4.43	4.10	3.87	3.70	3.56
21	8.02	5.78	4.87	4.37	4.04	3.81	3.64	3.51
22	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45
23	7.88	5.66	4.76	4.26	3.94	3.71	3.54	3.41
24	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36
25	7.77	5.57	4.68	4.18	3.85	3.63	3.46	3.32
30	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17
40	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99
60	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82
120	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66
∞	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51

续表Ⅶ　F分布表($\alpha=0.01$)

N_2 \ N_1	9	10	12	15	20	60	∞
1	6022	6056	6106	6157	6209	6313	6366
2	99.39	99.40	99.42	99.43	99.45	99.48	99.50
3	27.35	27.23	27.05	26.87	26.69	26.32	26.13
4	14.66	14.55	14.37	14.20	14.02	13.65	13.46
5	10.16	10.05	9.89	9.72	9.55	9.20	9.02
6	7.98	7.87	7.72	7.56	7.40	7.06	6.88
7	6.72	6.62	6.47	6.31	6.16	5.82	5.65
8	5.91	5.81	5.67	5.52	5.36	5.03	4.86
9	5.35	5.26	5.11	4.96	4.81	4.48	4.31
10	4.94	4.85	4.71	4.56	4.41	4.08	3.91
11	4.63	4.54	4.40	4.25	4.10	3.78	3.60
12	4.39	4.30	4.16	4.01	3.86	3.54	3.36
13	4.19	4.10	3.96	3.82	3.66	3.34	3.17
14	4.03	3.94	3.80	3.66	3.51	3.18	3.00
15	3.89	3.80	3.67	3.52	3.37	3.05	2.87
16	3.78	3.69	3.55	3.41	3.26	2.93	2.75
17	3.68	3.59	3.46	3.31	3.16	2.83	2.65
18	3.60	3.51	3.37	3.23	3.08	2.75	2.57
19	3.52	3.43	3.30	3.15	3.00	2.67	2.49
20	3.46	3.37	3.23	3.09	2.94	2.61	2.42
21	3.40	3.31	3.17	3.03	2.88	2.55	2.36
22	3.35	3.26	3.12	2.98	2.83	2.50	2.31
23	3.30	3.21	3.07	2.93	2.78	2.45	2.26
24	3.26	3.17	3.03	2.89	2.74	2.40	2.21
25	3.22	3.13	2.99	2.85	2.70	2.36	2.17
30	2.07	2.98	2.84	2.70	2.55	2.21	2.01
40	2.89	2.80	2.66	2.52	2.37	2.02	1.80
60	2.72	2.63	2.50	2.35	2.20	1.84	1.60
120	2.56	2.47	2.34	2.19	2.03	1.66	1.38
∞	2.41	2.32	2.18	2.04	1.88	1.47	1.00

表Ⅷ　相关系数检验表

$n-2$	5%	1%	$n-2$	5%	1%	$n-2$	5%	1%
1	0.997	1.000	16	0.468	0.590	35	0.325	0.418
2	0.950	0.990	17	0.456	0.575	40	0.304	0.393
3	0.878	0.959	18	0.444	0.561	45	0.288	0.372
4	0.811	0.917	19	0.433	0.549	50	0.273	0.354
5	0.754	0.874	20	0.423	0.537	60	0.250	0.325
6	0.707	0.834	21	0.413	0.526	70	0.232	0.302
7	0.666	0.798	22	0.404	0.515	80	0.217	0.283
8	0.632	0.765	23	0.396	0.505	90	0.205	0.267
9	0.602	0.735	24	0.388	0.496	100	0.195	0.254
10	0.576	0.708	25	0.381	0.487	125	0.174	0.228
11	0.553	0.684	26	0.374	0.478	150	0.159	0.208
12	0.532	0.661	27	0.367	0.470	200	0.138	0.181
13	0.514	0.641	28	0.361	0.463	300	0.113	0.148
14	0.497	0.623	29	0.355	0.456	400	0.098	0.128
15	0.482	0.606	30	0.349	0.449	1000	0.062	0.081

表Ⅸ　数字1到9的随机排列表

	1	2	3	4	5
	549716832	791642583	934516287	754329681	896723541
	513927648	436819527	683241597	469317258	496387152
0	623794815	783692451	714258963	913245687	643792851
	782463195	134879526	572813694	815294376	851736429
	129356748	527483196	527316498	726845913	681249573
	478562139	916427385	967358724	961543827	246718935
	395261874	215839764	923764851	741269835	148356927
1	318594267	218796453	971546328	168327594	983654172
	318246579	294836751	317842569	714289536	971256834
	724938651	985642731	862915473	148239576	789213465

	1	2	3	4	5
	892147365	296541873	941275386	953721846	287354146
	754639182	841726593	815932674	295374618	275841936
2	467593281	197524863	634172589	314862759	157938426
	681324975	751436982	148726935	829653714	895462317
	326597874	349561287	726138495	783914256	975641283
	983756421	285947136	536249817	723456981	536471892
	796812453	482679513	862371594	846179352	165789231
3	437586192	285749163	621837945	561837492	941276853
	967124853	186527394	146385729	382614579	278654913
	468932715	963854271	231486957	573124869	475938621
	917582634	278561943	127394565	572193846	532794618
	274139586	475938612	968537142	136982754	284957361
4	274938156	857934261	815924673	678245931	653782491
	827135649	345761892	321956847	467312958	238461975
	841329765	165798432	387964152	917632485	862153974
	247689531	256819347	385761429	754368912	145386297
	783592461	639572841	189267534	815934267	641823759
5	952164378	529387164	761394528	673512489	873695214
	378412596	478359162	723456981	148726953	517342869
	536928147	861543927	376194528	815793426	876521489
	836249157	893641572	783591264	298165437	864352971
	371956428	436958721	891782435	935216847	169532874
6	142879356	987423165	672198534	749853162	987346215
	952768143	429615387	543718926	375186249	478139652
	462891753	217945386	571826439	471283956	186459273
	827561934	584793126	967543281	431576829	839624751
	625489137	413569783	654973128	963845271	674158932
7	236419857	586214973	514983762	893521746	648259173
	239581467	836211759	239867415	217984365	538741692
	495321867	371459623	286495371	965421738	658741923

	1	2	3	4	5
	627548139	789125643	186352971	835694217	478621359
	534279186	456729318	647298153	475238196	318592764
8	379254618	294187653	452698137	916427835	394625817
	458913267	485672193	498126357	986327514	713249658
	892673145	279351684	617859243	152984763	937184256
	964723815	183654729	817956342	125864973	473596218
	517296438	231698754	614759832	173248965	874695312
9	372581694	293715846	562193487	782459613	623841395
	851962437	627451938	258649137	617594283	274386195
	249587136	235896171	298764351	136542879	687529134

表X　　200个随机数字表

12	67	78	24	44	54	12	73	97	48	79	91	20	20	17	31	83	20	85	66
06	24	89	57	11	27	43	03	14	29	84	52	86	13	51	70	65	88	60	88
29	15	84	77	17	86	64	87	06	55	36	44	92	58	64	91	94	48	64	65
49	56	97	93	91	59	41	21	98	03	70	95	31	99	74	45	67	94	47	70
50	77	60	28	58	75	70	96	70	07	60	66	05	95	58	39	20	25	96	89
00	31	32	48	23	12	31	08	51	06	23	44	26	43	56	34	78	65	50	80
01	67	45	57	55	98	93	69	07	81	62	35	22	03	89	22	54	94	83	31
24	00	48	34	15	45	34	50	02	37	43	57	36	13	76	71	95	40	34	10
77	52	60	27	64	16	06	83	38	73	51	32	62	85	24	58	54	29	64	56
36	29	93	93	10	00	51	34	81	26	13	53	26	29	16	94	19	01	40	45
94	82	03	96	49	78	32	61	17	78	70	12	91	69	99	62	75	16	50	69
23	12	21	19	67	27	86	47	43	25	25	05	76	17	50	55	70	32	83	36
77	58	90	38	66	57	45	85	13	93	00	65	30	59	39	44	86	75	90	73
92	37	51	97	83	78	12	70	41	42	01	72	10	48	88	95	05	24	44	21
28	93	48	44	13	02	49	32	07	95	26	47	67	70	72	71	08	47	26	18
09	68	01	98	80	27	49	78	56	67	40	22	13	66	61	33	53	18	36	03
61	73	92	33	89	48	20	42	32	33	79	37	68	88	44	59	35	17	97	61
82	35	37	33	53	42	52	04	14	54	08	25	48	89	57	87	59	89	96	76
39	20	77	72	55	19	66	58	57	91	38	43	67	97	52	66	45	39	74	67
51	90	71	05	82	38	37	40	94	52	24	09	35	44	37	33	35	20	65	84

97	49	53	79	17	25	02	65	77	70	88	45	53	51	63	30	89	66	42	03
73	18	91	38	25	82	29	71	56	89	86	74	68	58	75	36	93	13	33	31
17	79	34	97	25	89	01	17	67	92	62	25	54	70	52	88	28	05	61	17
97	27	26	86	17	67	59	56	95	07	49	05	70	06	70	35	23	35	26	18
56	06	63	00	07	40	65	87	09	49	70	31	67	02	33	39	04	40	01	51
43	83	39	24	50	74	10	05	38	11	25	80	44	14	98	21	87	41	02	74
63	19	91	27	08	59	02	28	47	13	05	53	02	28	81	91	46	90	95	52
23	87	60	31	98	87	76	57	82	47	64	87	50	45	73	54	26	47	62	10
07	04	47	34	36	03	87	67	03	28	72	19	98	99	32	98	78	76	85	40
98	61	67	62	09	89	73	50	06	81	29	09	43	43	30	12	32	69	52	19
36	86	50	21	42	18	20	55	00	90	01	96	42	12	68	18	45	93	52	99
70	64	92	95	09	09	27	63	09	29	69	99	98	26	19	83	94	88	95	37
41	71	91	61	31	86	38	01	71	79	44	75	67	69	35	31	69	47	81	64
23	48	32	36	88	50	29	07	27	32	21	28	73	41	77	39	00	78	92	65
13	32	97	81	00	28	87	13	00	86	56	16	81	20	63	29	37	45	08	91
70	55	85	27	24	96	91	83	89	17	89	98	51	31	17	29	05	77	62	95
12	50	54	01	63	40	74	86	88	90	63	76	97	74	08	70	88	88	98	96
97	99	24	63	47	63	47	66	21	79	28	66	67	24	33	20	01	52	09	59
16	99	63	89	67	89	14	55	70	31	45	56	05	71	84	30	48	32	90	94
57	95	93	54	30	74	11	18	31	26	75	39	81	28	63	34	31	23	77	67
01	32	91	11	23	65	44	58	69	77	58	80	35	20	92	12	48	15	56	67
00	30	26	68	89	38	13	99	47	38	06	82	49	47	40	33	23	72	01	50
48	15	27	13	97	70	18	48	14	28	26	30	74	16	13	07	36	21	94	84
58	86	65	76	67	05	99	53	33	56	92	61	63	98	55	39	15	77	61	67
75	07	14	81	41	16	12	21	79	82	16	42	70	43	73	33	78	22	63	25
86	19	97	09	64	04	21	26	65	11	20	32	82	38	52	94	79	21	85	07
66	17	52	10	35	14	21	89	54	32	61	49	63	06	36	25	63	84	78	24
56	70	95	77	25	19	21	15	29	88	57	75	51	19	31	06	48	50	09	65
14	43	67	32	81	78	19	72	32	70	34	86	11	90	37	02	54	39	45	87
04	17	91	71	96	90	85	68	32	35	77	20	71	43	55	95	28	90	51	69

后 记

本教材由王铭文主编，仲崇骥参加编写第三、四章的教材本文、学习指导及习题解答，朱显海参加编写第五、六章的教材本文、学习指导及习题解答，李淑云、徐云波二同志协助复查了习题的解答。

本教材是大学数学系自学丛书之一。它的内容符合目前高等师范院校《概率论与数理统计》教学大纲的要求。

为了便于自学，本教材各章除本文外都编有学习指导及习题解答。在学习指导中总结了各章的内容要点，对在自学中不易理解或容易误解的一些内容，通过举例作了较详细的说明，并对有些重要内容作了系统归纳或进一步引伸。

在编写过程中得到东北师范大学数学系及概率统计教研室的热情支持，又经辽宁大学数学系刘良和副教授协助审校，并提出很好的修改意见，谨此致谢。

由于我们水平所限，虽然力求写好，仍可能有不妥与错误之处，衷心希望读者给予批评指正。

编　者

1982.3

［Ｇｅｎｅｒａｌ Ｉｎｆｏｒｍａｔｉｏｎ］
书名＝概率论与数理统计
作者＝
页数＝５１２
ＳＳ号＝０
出版日期＝
Ｖｓｓ号＝８８６９６７０１